THE UNIVERSITY
of EDINBURGH

Mathematical modelling of photoperiodic external coincidence mechanisms in the model plant, *Arabidopsis thaliana*

Robert W. Smith

Doctor of Philosophy
University of Edinburgh
2013

Declaration

I declare that I, Robert W. Smith, have composed this thesis and that I have made a substantial contribution to the presented work under the supervision of Prof. Andrew Millar and Dr. Karen Halliday. Work contributed by others (collaborators or members of the research groups that I belong to) in this thesis is indicated clearly at the relevant points. This work has not been submitted for any other degree or professional qualification.

The techniques and results presented in this thesis have been published in:

Y. H. Song, R. W. Smith, B. J. To, A. J. Millar, T. Imaizumi (2012) FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science*, vol. 336, pp. 1045-9.

J. Keily, D. R. MacGregor, R. W. Smith, A. J. Millar, K. J. Halliday, S. Penfield (2013) Model selection reveals control of cold signalling by evening-phased components of the plant circadian clock. *The Plant Journal*, vol. 76, pp. 247-57.

Further submissions include:

R. W. Smith, D. D. Seaton, Y. H. Song, D. R. MacGregor, K. Stewart, G. Steel, J. Foreman, S. Penfield, T. Imaizumi, A. J. Millar, K. J. Halliday. Linked circadian outputs control growth and development in response to photoperiod and temperature.

Y. H. Chew, R. W. Smith, H. J. Jones, D. D. Seaton, R. Grima, K. J. Halliday. Mathematical models light up plant signalling.

Signed: Date:

Acknowledgements

The first people I wish to acknowledge are my supervisors, Prof. Andrew Millar and Dr. Karen Halliday, who have provided me with excellent support and advice throughout my PhD studies. The opportunities that they have made available to me, such as attending and presenting at national & international conferences, being part of a number of collaborative projects, and taking part in outreach activities, have given me some invaluable experiences that I hope will aid me as I move forward with my career.

At this point I would also like to thank all of the current and past members of both the Millar and Halliday research groups that have worked at the University of Edinburgh during my studies. Notably Drs. Alexandra Pokhilko and Daniel Seaton for their advice and collaborations on mathematical projects; Drs. Julia Foreman, Jayne Griffiths, Gabriela Toledo-Ortiz, Kelly Stewart and Keun-Pyo Lee as well as Gavin Steel who have provided vital experimental support that has contributed to this thesis; Dr. Yin Hoon Chew, Harriet Jones, and Dr. Ramon Grima, who I have collaborated with on a review paper; Dr. Tomasz Zielinski for computational/BioDare support, and; all those that have baked cakes for the office (particularly Dr. Anne Moore). Furthermore, I would like to thank all current and past members of the BBSRC-funded ROBuST project. Notably Drs. Steven Penfield, Dana MacGregor, Kate Sidaway-Lee and Jack Keily from the University of Exeter with whom I have collaborated with on projects contained in this thesis. From international collaborations, I am very grateful to Asst. Prof. Takato Imaizumi and Dr. Young Hun Song with whom I have collaborated with on a number of projects within this thesis. Without their high quality data the development of the flowering models in Chapters 3 & 4 could never have started.

From my time at Warwick University, studying for an MSc in Systems Biology, I wish to acknowledge those that I started my PhD years alongside. In particular, future Drs. Kate Richardson, Phil Law, Jo Hulsmans and Ben Wareham who I have shared numerous wide-ranging discussions (normally starting with science) and drinks with. Also I would like to thank those at the Systems Biology Doctoral Training Centre at Warwick University, notably Profs. David Rand & Vicky Buchanan-Wollaston and Drs. Katherine Denby, Miriam Gifford & Mirela Domijan. During my time at Edinburgh, I would like to thank the mates that have provided me with light relief during my studies. In particular, Tom Southgate, Tim Gallant, Greg Russell, Pete Sanderson, Sebs Borggrewe, "Dutch" Tom Beving and all those at Edinburgh Medics and Waverley Inveresk Trinity hockey clubs.

Finally, I would like to say a massive thank you to my parents Nigel & Barbara, my younger brother Pete, and the "extended family" George Neal & Matt West. You guys have provided me with a huge amount of support since I started my undergraduate studies at Newcastle and I am very grateful for it!

The work contained in this thesis is supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC) grants BB/F59011/1 and BB/F005237/1.

Abstract

As plants are sessile organisms, processes controlling plant growth and development must react to fluctuations in the external environment to aid plant survival. However, as the climate of the Earth changes and becomes more extreme, plants become less able to develop to their optimal capacity and this can have an adverse effect on crop yield and biofuel feedstock production. Thus, it is becoming increasingly important to understand the molecular mechanisms used by plants to respond to external stimuli. One important system that plants utilise in their response to environmental fluctuations is the circadian clock. The circadian clock is a time-measuring device that buffers the timing of plant growth and development against fluctuations in the local environment, such as temperature, light quality and light intensity. Importantly, the circadian clock is also able to measure day-length (photoperiod). Thus, plant development and growth is co-ordinated with photoperiod that is closely linked to seasonal changes. A key example of this is the time taken for a plant to flower. Flowering of *Arabidopsis thaliana* occurs specifically in long-days (LDs) of spring/summer months. Thus, the circadian clock is a key regulator promoting flowering in LD conditions.

In conjunction with experimental studies, mathematical modelling has proven to be a successful method of elucidating the mechanisms that underlie complex biological systems. One example of this 'systems biology' approach is in uncovering the components that make up the Arabidopsis circadian clock mechanism. Previous research in our group has also led to the development of a model describing photoperiodic flowering that is tentatively linked to the circadian clock mechanism. In this thesis I shall develop on these models to highlight five key results:

1. using rhythmic *PHYTOCHROME INTERACTING FACTOR 4* (*PIF4)* and *PIF5* mRNA as an example, I shall show that multiple circadian regulators

are required to describe rhythmic transcription of target genes across multiple photoperiods;

2. the stabilisation of CONSTANS (CO) protein by the blue light-signalling component FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1) is required to for flowering in LDs and has a relatively larger impact on photoperiodic flowering than FKF1-dependent degradation of CYCLING DOF FACTOR 1 (CDF1), an inhibitor of flowering;
3. multiple components of the circadian clock play specific post-translational roles in photoperiodic flowering to promote the acceleration of flowering specifically in LDs;
4. temperature regulation of photoperiodic flowering can be explained through an interaction between CO and PIF proteins, limiting the effects of temperature to a specific time-window in a 24-hour day;
5. red light- and temperature-control of the circadian clock can be explained by altering the post-translational regulation of circadian clock components.

Lay Summary

Plants are not afforded the luxury of mammals that are able to move to new locations as the external environment changes around them. Therefore, plants have to be able to deal with a wide range of environments flexibly and robustly so that energy directed towards plant growth and development is utilised in favourable seasonal conditions. However, as climates become more extreme this can lead to a lower crop yield and biofuel feedstock production that we require for food and energy. Thus, it is important to fully understand how plants react and adapt to fluctuations in their local growth environments. To this end, the plant *Arabidopsis thaliana* is used as a simple model of crop species. In combination with biological experiments, mathematical modelling has proven to be a useful method of describing complex biological systems and generating hypotheses to be tested in future experiments. Over the course of this thesis I shall use mathematical modelling to explore the molecular mechanisms that are utilised by Arabidopsis to adapt to changes in day-length, light quality and temperature. By building mathematical models, I will show that components of environmental signalling must act at multiple points within the molecular pathways governing plant development. Such networks are important in providing the flexibility and robustness required to prevent adverse development in unfavourable climates. Hence, by furthering our understanding of how environmental signals regulate the molecular mechanisms of Arabidopsis, this research can be used to help optimise crop development in ever-changing climates.

List of Figures

List of Tables

List of Abbreviations

ACT2	ACTIN2
ACT7	ACTIN7
ADO	ADAGIO
AIC	Akiake Information Criterion
AICc	corrected Akaike Information Criterion
$AICc^U$	corrected Akaike Information Criterion with Uncertainty
AP1	APETALA1
ASK	Arabidopsis Skp1-like
BBSRC	Biotechnology & Biological Sciences Research Council
bHLH	basic Helix-Loop-Helix
BIC	Bayesian Information Criterion
CAB2	CHLORPHYLL A/B BINDING PROTEIN 2
CBF	C-REPEAT BINDING FACTOR
CBS	CCA1 Binding Site
CCA1	CIRCADIAN CLOCK ASSOCIATED 1
CCT	CONSTANS, CO-like and TOC1
CDF	CYCLING DOF FACTOR
cdf1-R	CDF1 RNAi transgene
CHE	CCA1 HIKING EXPEDITION
ChIP	Chromatin Immunoprecipitation
CO	CONSTANS, CO-like and TOC1
Col	Columbia
COP1	CONSTITUTIVELY PHOTOMORPHOGENIC 1
COR	COLD REGULATED
cry	cryptochromes
Cu	CCA1 up-regulation
CUL	CULLIN
EC	Evening Complex

EE	Evening Element
ELF	EARLY FLOWERING
FBH	FLOWERING BHLH
FFT-NLLS	Fast Fourier Transform - Non-Linear Least Squares
FKF1	FLAVIN-BINDING, KELCH REPEATS, F-BOX 1
FT	FLOWERING LOCUS T
GI	GIGANTEA
H:C	Hot:Cold
HOS	HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE 1
HUD	Hormone Up at Dawn
HY5	LONG HYPOCOTYL 5
ICE	INDUCER OF CBF EXPRESSION
L:D	Light:Dark
LBS	LUX Binding Site
LD	Long Day
L*er*	Landsberg *erecta*
LHY	LATE ELONGATED HYPOCOTYL
LKP2	LOV KELCH PROTEIN 2
LOV	LIGHT-OXYGEN-VOLTAGE
LOV1	LONG VEGETATIVE PHASE 1
LUX	LUX ARRHYTHMO
ME	Morning Element
NP	NIGHT-TIME PEAK
ODE	Ordinary Differential Equation
ox	overexpressor
p(x)	probability of x
PAS	PER-ARNT-SIM
phy	phytochrome
PIF	PHYTOCHROME INTERACTING FACTOR
PRR	PSEUDO RESPONSE REGULATOR
PTO	Post-Translational Oscillator
(q)RT-PCR	(quantitative) Real Time Polymerase Chain Reaction

ROBuST	Regulation Of Biological Signalling by Temperature
ROS	Reactive Oxygen Species
RVE	REVEILLE
SAM	Shoot Apical Meristem
SASSy	Sensitivity Analysis Software for Systems
SBSI	Systems Biology Software Infrastructure
SCF	Skp/CULLIN/F-box
SD	Short Day
SPA	SUPPRESSOR OF PHYA
SUC2	SUCROSE TRANSPORTER 2
Td	TOC1 down-regulated
Td:Cu	TOC1 down-regulated : CCA1 up-regulated
Td:Xd	TOC1 down-regulated : X down-regulated
TI	Thermodynamic Integration
TIC	TIME FOR COFFEE
TI^P	Thermodynamic Integration with Penalty
TOC1	TIMING OF CAB EXPRESSION 1
TPL	TOPLESS
TTFL	Transcription-Translation Feedback Loop
UBQ10	UBIQUITIN 10
UVR	UV RESISTANCE LOCUR
WT	Wild-Type
ZT	Zeitgeber Time
ZTL	ZEITLUPE

Table of Contents

Chapter 1: Introduction

As climate change takes place, the fluctuations of environmental conditions leads to altered crop yields and biofuel feedstock production. Consequently, climate change has an impact on human and animal life that requires crops for food and energy sources. Recent studies have shown that climate change over the last 50 years has led to increased temperatures, decreased average solar radiation, an earlier onset of spring and a delayed start of autumn seasons [1, 2]. Using these environmental cues (as well as increased atmospheric CO_2) as markers of climate change, crop yields of 2050 and 2100 have been predicted to change in a latitude-dependent manner [3]. These responses can be partially reversed by changing the date at which crops are planted. Hence, the effects of climate change on crop yield are regulated by a complex interaction of environmental signals. To decipher how these environmental cues interact, the Biotechnology & Biological Sciences Research Council (BBSRC)-funded ROBuST (Regulation of biological signalling by temperature) project was created to determine how growth and development of the model plant, *Arabidopsis thaliana*, is mediated by light-controlled phytochromes (phys), temperature-regulated C-REPEAT BINDING FACTORs (CBFs) and the endogenous circadian clock that measures day-length. Each of these processes shall be introduced in more detail throughout this chapter.

In this thesis I shall concentrate on how flowering of Arabidopsis is regulated by the duration of light in a day (photoperiod), different qualities of light (red and blue) and warm ambient temperatures (~27°C). Through the use of mathematical modelling I shall be able to: first, describe new and published data, and; second, generate novel, experimentally testable hypotheses that have not previously been considered. A key advantage of mathematical modelling is that an entire biological network can be examined efficiently and cost-effectively. As shall be discussed in preceding chapters, when building mathematical models a balance needs to be struck between finding a simple representation of a complex network, whilst maintaining how well

the model is able to describe, or 'fit', experimental data. The construction of a 'good' mathematical model is constrained by the availability and quality of data. It is important that there is enough 'training data' (data used in model construction) to constrain model parameters and that this data is reproducible across numerous experiments in similar conditions. If the data is not reproducible then the resulting model may describe anomalies seen in single biological replicates. When a model simulation is unable to accurately match experimental data, then the model is not a good representation of the biological network. However, if a model is too complex, then it may 'overfit' the data. This implies that whilst the model simulation accurately describes the required data, the network is too large to systematically probe in the search for novel model-derived hypotheses. Whilst the development of models is constrained by the availability of data, once constructed a model is able to help direct experimental research by predicting the outcomes of experiments and generating hypotheses for future experiments.

In the rest of this chapter I shall introduce the biological networks that shall be of interest throughout this thesis, what data is currently available and, in some cases, how mathematical modelling has already helped develop our understanding of the biological mechanisms.

Regulation of plant development by day-length

The Circadian Clock

In order for plants to respond to changing photoperiods, molecular mechanisms must be able to measure day-length. One such mechanism is the circadian clock. Circadian (circa = 'about'; dian = 'a day') clocks are present in most free-living species to entrain internal biochemical processes to diurnal cycles (e.g. light:dark, L:D, or hot:cold, H:C) in the environment (for a review of recent research of plant and mammalian circadian clocks, please see [4]). Once entrained, the circadian clock then continues to regulate daily internal processes in the absence of cues from the external environment, such as dawn and dusk. One example of this is the 24hr (hour)

leaf movement rhythm of Arabidopsis plants that are transferred to constant light after being grown in 12L:12D (12hrs light: 12hrs dark) entraining conditions [5]. The length of the rhythms seen in constant environmental conditions is referred to as the period length – so, in this example, leaf movement of wild-type (WT) Arabidopsis has a 24hr period. Another key property of circadian clocks is that they are 'temperature compensated', such that the period of circadian rhythms is maintained at ~24hrs over an ambient temperature range (~12-27°C) despite temperature-induced changes in chemical reaction rates [6–8]. A similar form of compensation has been observed over ranges of light intensities in different light qualities [9, 10]. Therefore, the circadian clock synchronises the timing of biochemical processes in Arabidopsis with the external environment despite fluctuations in light- and temperature-signals. A circadian clock that is synchronised with the environment helps improve plant fitness [11].

A number of studies have used some of the key circadian properties, namely the 24hr period length in constant environmental conditions, to elucidate components that are part of the circadian clock mechanism. By using experimental observations alongside mathematical modelling, the Arabidopsis circadian clock is proposed to feature a number of feedback loops that, at its core, comprise a three-component repressilator mechanism (see Figure 1.1; [12]). Theoretical studies of oscillating systems have proposed that such a structure of the circadian clock would provide the flexibility and robustness required for 24hr rhythms to be maintained despite fluctuations in the external environment [13, 14]. The first components of the repressilator are *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) & *LATE ELONGATED HYPOCOTYL* (*LHY*), whose transcription peaks at around dawn in diurnal cycles [15–17]. Removal of these components leads to short period circadian rhythms and a phase advance (earlier peak expression) of circadian regulated transcripts in diurnal cycles [17, 18]. These components encode two homologous MYB transcription factors that can act as transcriptional regulators of target genes through the evening element (EE, AAATATCT) and CCA1 binding site (CBS, AAAAATCT) promoter motifs (see Chapter 2; [19]). Importantly, CCA1 & LHY represses the transcription of genes active from dusk to dawn, in particular *TIMING OF CAB EXPRESSION 1*

(*TOC1*), *GIGANTEA* (*GI*) and components of the Evening Complex (EC): *EARLY FLOWERING 4* (*ELF4*) – *ELF3* – *LUX ARRHYTHMO* (Figure 1.1; [20–25]). The ELF4 protein aids ELF3 localisation to nuclear speckles, where ELF3 interacts with the DNA-binding MYB protein LUX to form the EC [25, 26]. The EC goes on to inhibit transcription of target genes, namely *PRR9* and *LUX* through the LUX binding site (LBS, GATWCG), whilst the clock model also includes EC inhibition of *TOC1*, *GI* and *ELF4* due to the presence of LBS motifs in their promoters and increased expression in *elf3* loss-of-function transgenic plants (see Chapter 2; [12, 25, 27, 28]). Importantly, the PRR protein products, including PRR5, act as transcriptional repressors of *CCA1* & *LHY* completing the CCA1/LHY-EC-PRRs repressilator mechanism (Figure 1.1; [29–31]). Further to this repressilator mechanism, GI-mediated protein interactions that may sequester ELF3 & ELF4, as well as play a role in TOC1 degradation through the blue light-regulated GI-ZEITLUPE (ZTL) complex formation, fine-tune the circadian clock mechanism (Figure 1.1; [32–35]). As we shall see later, post-translational interactions by GI are important for the circadian regulation of flowering in Arabidopsis.

The repressilator model of the Arabidopsis circadian clock has evolved over the course of six publications that have used experimental data to construct mathematical models that are able to describe the circadian clock and hypothesise missing components of the system [12, 23, 36–39]. These models have been built using ordinary differential equations (ODEs) that measure the change of relative concentration levels over time. The use of ODEs is reflected by the type of data that is available. In this case, time-series data of mRNA and protein levels are obtained from whole seedlings (i.e. a population of cells) rather than from single cells. The ODE model therefore represents effects seen from a population of circadian clocks. An example ODE equation in this context takes the form

$$\frac{dc}{dt} = p - kc \tag{1.1}$$

where c is the concentration of interest, p is the production rate and k is the degradation rate that, in these models, is proportional to the concentration c. The

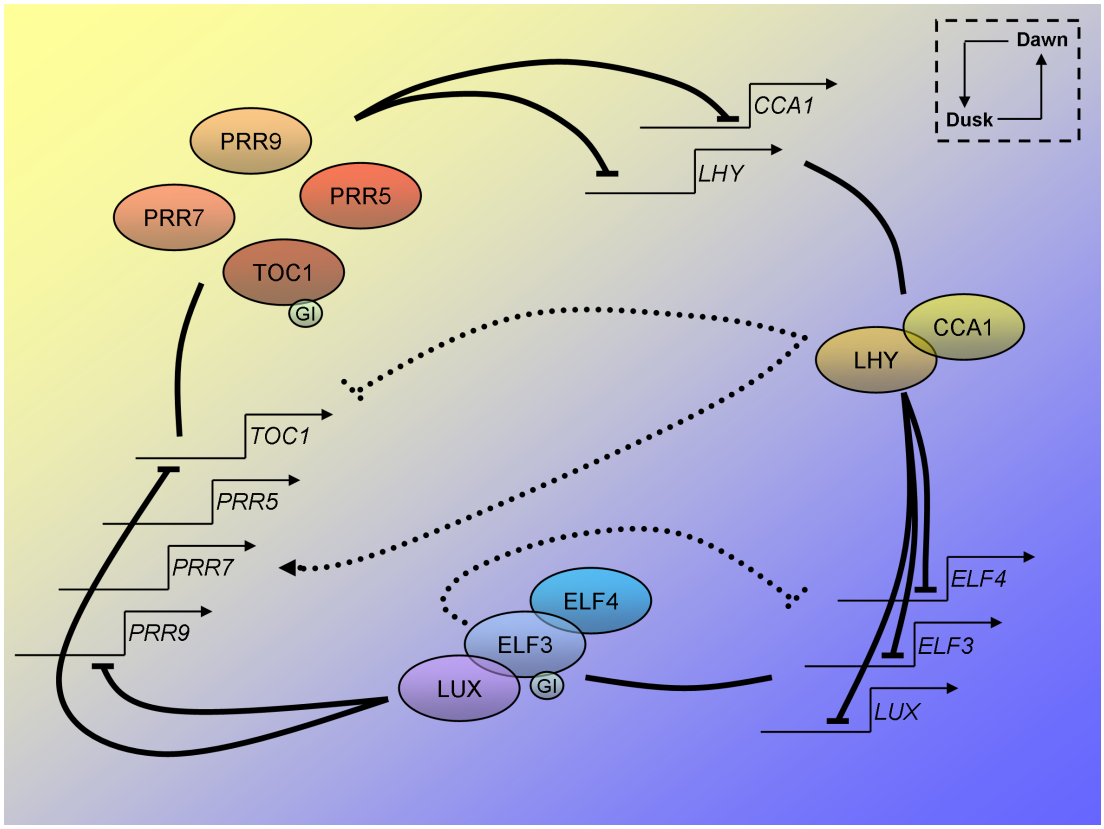


Figure 1.1: Schematic of circadian clock repressilator. Transcription of *CCA1* & *LHY* is activated at dawn. The proteins go on to inhibit transcription of the key night-time components *ELF4*, *ELF3* and *LUX* that form the evening complex (EC). The EC then inhibits the transcription of *PRR9* and *TOC1* who, along with other PRR proteins, inhibit the transcription of *CCA1/LHY* throughout the day. This mechanism is highlighted in the solid black lines. The dashed black lines represent additional connections within the circadian clock mechanism. Notably, the EC can feed back on the transcription of its own components, *ELF4* and *LUX*, whilst CCA1 & LHY proteins also regulate transcription of the *PRR*s. Furthermore, post-translational modifications of ELF3 and TOC1 by GI also play an important role in the circadian clock model.

parameter values (that represent transcription, translation & degradation rates and binding/interaction efficiencies of proteins, such as p and k in (1.1)) are constrained by high resolution time-series data where expression levels are generally measured every 2-4hrs. The term $\mathrm{d}c/\mathrm{d}t$ represents the change in c ($c(t+\Delta t) - c(t)$) over a small time-step, Δt. This simple equation can be solved analytically such that

$$c(t) = \left(c_0 - \frac{p}{k}\right)e^{-kt} + \frac{p}{k}. \tag{1.2}$$

If p, k are positive constants and the initial condition of c, $c(t = 0) = c_0 \geq 0$ then $c \geq 0$ for all t. However, in large systems of coupled ODE equations, as with circadian clock models, it is very hard to solve the system analytically. Therefore, the system is solved numerically by updating the equations with the concentrations at the preceding time-point. Whilst this system is built to describe a population of cells, ODEs can be used to describe single cell dynamics in cases where components of the modelled system have large copy numbers relative to other molecules within the cell. Although the circadian clock has not been extensively analysed experimentally in single cells, theoretical analysis that scaled concentrations of the 'Pokhilko2010' circadian clock model, that did not include EC dynamics (see below), to describe the case where clock components have low copy numbers compared to other molecules in the cell was able to explain why circadian rhythms dampen in constant light through stochastic fluctuations in concentrations of clock components [38, 40]. As all the time-series data presented in this thesis shall be obtained from whole seedlings, the models that are constructed will use ODE systems describing effects that are averaged over a population of cells.

Importantly, the circadian clock models have allowed the Millar group (University of Edinburgh, UK) to propose roles for components whose function in the clock mechanism had not yet been fully elucidated. For example, the 'Locke2005' clock model featured two inter-locked feedback loops made up of 4 components (see Chapter 2; [23]). The morning genes, CCA1/LHY, inhibited transcription of *TOC1* and component *Y*. These components formed a second feedback loop whereby Y activated *TOC1* transcription and TOC1 repressed *Y* mRNA. The evening and morning components are then connected through TOC1-mediated activation of *CCA1/LHY* mRNA using component *X* to delay the activating signal [23]. Hence, the model proposed two hidden components, X and Y. Component Y was proposed to play the role of GI in the system and this was, in part, validated experimentally [23, 36]. The identity of component X was suggested to be ELF4 based on simulations of transcriptional profiles [23]. It was not until the 'Pokhilko2011' clock model was created that the need for this hypothetical component was obviated by changing TOC1 regulation of *CCA1/LHY* mRNA from activation to inhibition and by

introducing the EC [12]. Model variants created between the 'Locke2005' and 'Pokhilko2011' models incorporated the roles of PRR components forming a feedback loop with CCA1/LHY and the transcriptional inhibition of *PRR9* by TOC1 [36–38]. The most recent model variant, the 'Pokhilko2012' model, now incorporates all of the transcriptional targets of TOC1 within the clock system and is able to produce low amplitude rhythms of *CCA1/LHY* in both *toc1* loss-of-function mutants and *TOC1* overexpressing (*TOC1-ox*) transgenic lines, a detail that had previously been unable to be explained, as well as induction of *TOC1* transcription by abscisic acid (ABA) [39].

The 'Pokhilko2012' model is able to accurately describe a wide range of behaviours observed experimentally without requiring the need for hidden components (such as X and Y). However, a number of genes have been shown experimentally to form part of the circadian clock mechanism through genetic perturbation. Here, I shall introduce four groups of genes have been found to play a role in the circadian clock that are not currently part of the mathematical model. First, the TIME FOR COFFEE (TIC) nuclear protein, whose transcription is not regulated by the circadian clock, has been shown to alter circadian rhythms in diurnal conditions, period lengths in constant light, and the input of light signals into the clock [41, 42]. Interestingly, TIC has been shown to interact with two basic helix-loop-helix (bHLH) proteins, MYC2 and PHYTOCHROME INTERACTING FACTOR 5 (PIF5), which are important in jasmonate and light signalling (see below), respectively [43]. Thus, it may be that TIC regulates the clock specifically through these protein interactions mediating external signals into the circadian clock system. Furthermore, a group of circadian regulated MYB transcription factors homologous to CCA1 & LHY called REVEILLE 4 (RVE4), RVE6 and RVE8 target EE promoter motifs to activate transcription of evening phased genes [44–46]. Thus, it seems that CCA1 & LHY proteins inhibit these targets at the start of the day and they are then activated by RVE proteins at dusk. A recent report has also implicated further RVE family members in the regulation of the circadian clock [47]. Another gene that forms a new feedback loop within the circadian clock is CCA1 HIKING EXPEDITION (CHE) that interacts with TOC1 to inhibit *CCA1* mRNA [48]. Transcription of *CHE* is

inhibited by CCA1 & LHY, completing the feedback regulation. Finally, a recent study has implicated genes induced by light during the dark period of L:D cycles that also regulate the clock [49]. The *LNK1* and *LNK2* transcripts are regulated by the PRR proteins, while the *lnk1;2* loss-of-function mutant has remarkable similarities to the *prr9;7* loss-of-function transgenic plant [49, 50]. This suggests a potential mechanism whereby PRR9 and PRR7 may require LNK1 and LNK2, who have no recognised protein domains, to function properly. Recent studies have shown a similar role for TOPLESS (TPL) [51]. Whilst these and further components are clearly part of the biological circadian clock mechanism, it may be that their role is redundant with components already described in the 'Pokhilko2012' system. Thus, adding these to the mathematical model would make the system a lot more complex but may not provide biologists with much new information about the circadian system. The balance of model complexity with biological realism has been described at the start of this chapter and will be described in more detail in Chapter 2.

Circadian clock models of other species

Mathematical models have also helped elucidate knowledge of circadian clocks in other species. Here, the circadian clock mechanisms of mammals and two plant ancestors will be briefly introduced with comparisons drawn to the Arabidopsis system.

A simple view of the mammalian circadian clock, as observed in mice, is that it consists of three protein complexes and two negative feedback loops [52]. The first protein complex is a dimer of two bHLH/PER-ARNT-SIM (PAS) domain proteins BMAL1 and CLOCK. This protein complex activates transcription through binding to E-box motifs (CANNTG) in 'bursts' due to a continual turnover of complexes bound to the promoters [53, 54]. Specifically, BMAL1/CLOCK complexes activate transcription of *Rev-erb* genes whose proteins dimerise to repress transcription of *Bmal1*. The second feedback loop consists of activation of *Per* and *Cry* genes transcription by the BMAL1/CLOCK complex. The PER and CRY proteins then dimerise and inhibit the function of the BMAL1/CLOCK complex preventing further

activation of transcription. Whilst this clock mechanism is not as complex as that of the Arabidopsis clock model described above, experimental analysis has been able to closely examine protein dynamics of the mammalian circadian system. Experimental studies have observed that circadian rhythms are dependent on the ratio of BMAL1/CLOCK:PER/CRY protein complexes and that the ratio of these two complexes alters the robustness of circadian rhythms [55]. Mathematical analysis supported this conclusion and showed that circadian rhythms are most robust when there is a 1:1 ratio between the positive and negative complexes [52]. Further modelling highlighted that a second negative regulator of BMAL1/CLOCK complex function, hypothesised to be REV-ERB proteins, provides a clock mechanism whose rhythms are less sensitive to the ratio of complexes. Thus, rhythms oscillate with a period maintained at 24 hours over a wide range of protein stoichiometries, as is the case in different cell types in mammals [52].

Mathematical modelling of circadian clocks has also been used in the alga *Ostreococcus tauri* and the cyanobacteria *Synechococcus elongatis*. The circadian clock of Ostreococcus is an early ancestor of the Arabidopsis clock, featuring a negative feedback loop between *OtCCA1* and *OtTOC1*. Mathematical modelling has shown that, provided the model has enough light inputs, this simple system can support 24hr rhythms across a range of photoperiods without the extra complexity found in the Arabidopsis models [56, 57]. Interestingly, the circadian clock of cyanobacteria, another ancestor of plants, can be reconstituted *in vitro* such that rhythms can be observed when the correct components are added to a test-tube. This mechanism forms an example of a post-translational oscillator (PTO), unlike the transcription-translation feedback loops (TTFLs) in Arabidopsis, mammals and Ostreococcus. In this system three proteins KaiA, KaiB and KaiC interact to form 24hr rhythms of KaiC phosphorylation status [58]. Mathematical modelling of this simple system showed that these three components and phosphorylation alone are able to support 24hr oscillations [58]. Recently, *in vivo* cyanobacteria rhythms of the PTO have been shown to be intricately linked to a TTFL mechanism to coordinate the PTO with the onset of light [59].

These examples highlight three areas of circadian clock research that would be advantageous to explore within Arabidopsis. First, the mammalian circadian clock has highlighted the importance of balancing the ratio of protein complexes within the clock system. As the Arabidopsis mechanism can be simplified to a repressilator of three protein complexes, this analysis would yield new insights into how the plant circadian clock operates. Second, the key feature of the *in vitro* PTO oscillator of cyanobacteria is phosphorylation of KaiC. Currently, knowledge of phosphorylation of clock proteins in Arabidopsis is limited and whether 24hr cycles of phosphorylation of core components is important in the circadian system is unknown. Third, the *in vivo* linking of circadian TTFL and PTO oscillators in cyanobacteria suggests that the circadian clock mechanism in Arabidopsis may become more complex as a PTO is uncovered. As yet, the PTO of Arabidopsis has not been elucidated, although an evolutionary conserved marker for a 24hr PTO has been recently discovered [60].

Circadian regulation of plant growth & development

As mentioned previously, one marker for circadian function is rhythmic leaf movement of Arabidopsis [5]. This suggests that the circadian clock is able to regulate plant physiology. Observations have shown that hypocotyl elongation, photosynthesis, starch turnover, stomatal opening, cold acclimation and a number of hormones in Arabidopsis are regulated by the circadian clock [11, 20, 61–65]. Similar control of development by the circadian clock is seen in other species, such as sleep-wake cycles in mammals [66]. Importantly, a number of the components required for the regulation of Arabidopsis development in response to environmental changes have homologs in rice, *Oryza sativa*, and legume, *Lotus japonicus* [67–71]. Thus, the observations that are made experimentally in *Arabidopsis thaliana* can be translated into more complex plants, including crops. In this thesis I shall concentrate on how the circadian clock is able to regulate flowering in Arabidopsis.

Plants are often characterised by the day-lengths that accelerate the flowering process. For example, *Arabidopsis thaliana* flowers faster when grown under long

photoperiods and is, thus, referred to as a long-day (LD) plant [72]. The LD specific acceleration of flowering in Arabidopsis is regulated by *FLOWERING LOCUS T* (*FT*) mRNA levels. In LDs, *FT* transcription reaches maximal expression at dusk, whilst in short days (SDs) *FT* remains at minimal levels (Figure 1.2; [73–75]). Furthermore, loss-of-function *ft* transgenic plants are unable to accelerate flowering in LDs. Transcription of *FT* is regulated by the circadian clock such that, under constant light, *FT* mRNA is rhythmic with maximal expression during subjective night [20]. Under diurnal cycles, *FT* expression peaks at dusk and perturbing levels of clock components alters this rhythm [50, 76–78]. In addition, the positive correlation between photoperiod and *FT* transcription leads to increased *FT* mRNA levels observed in LDs compared to SDs (see Figure 1.2). This photoperiodic regulation represents an example of external coincidence [79]. External coincidence takes place when oscillations generated by the circadian clock occur during a specific time window of environmental diurnal cycles. Importantly, components of the circadian clock maintain a similar phase across multiple photoperiods [80]. Thus, the coordination of a circadian activator of *FT* transcription with the second half of the light period in LDs leads to the acceleration of flowering seen experimentally.

The pathway that connects the circadian clock to *FT* transcription has been elucidated experimentally. The key regulator of *FT* expression is CONSTANS (CO) such that, in *co* loss-of-function mutants, *FT* mRNA remains at basal levels in LDs leading to delayed flowering [73]. Transcription of *CO*, like *FT*, is under the control of the circadian clock, showing rhythms in constant light and diurnal cycles that are perturbed by loss-of-function or overexpression of clock components [50, 73, 76–78]. Furthermore, expression rhythms of *CO* mRNA are photoperiod-dependent. In SDs, *CO* mRNA has a single peak of expression during the night. However, in LDs, *CO* transcription has a double peak of expression (Figure 1.2; [34, 50, 73, 74, 76, 81]). The first of these LD expression peaks occurs prior to dusk and is dependent on the presence of FLAVIN-BINDING, KELCH REPEATS, F-BOX 1 (FKF1), a member of the ZTL protein family, that forms a blue light-dependent complex with GI to activate *CO* transcription [74, 81]. Protein levels of GI and FKF1 are regulated by the circadian clock with peak expression approximately 8-10 hours after dawn in

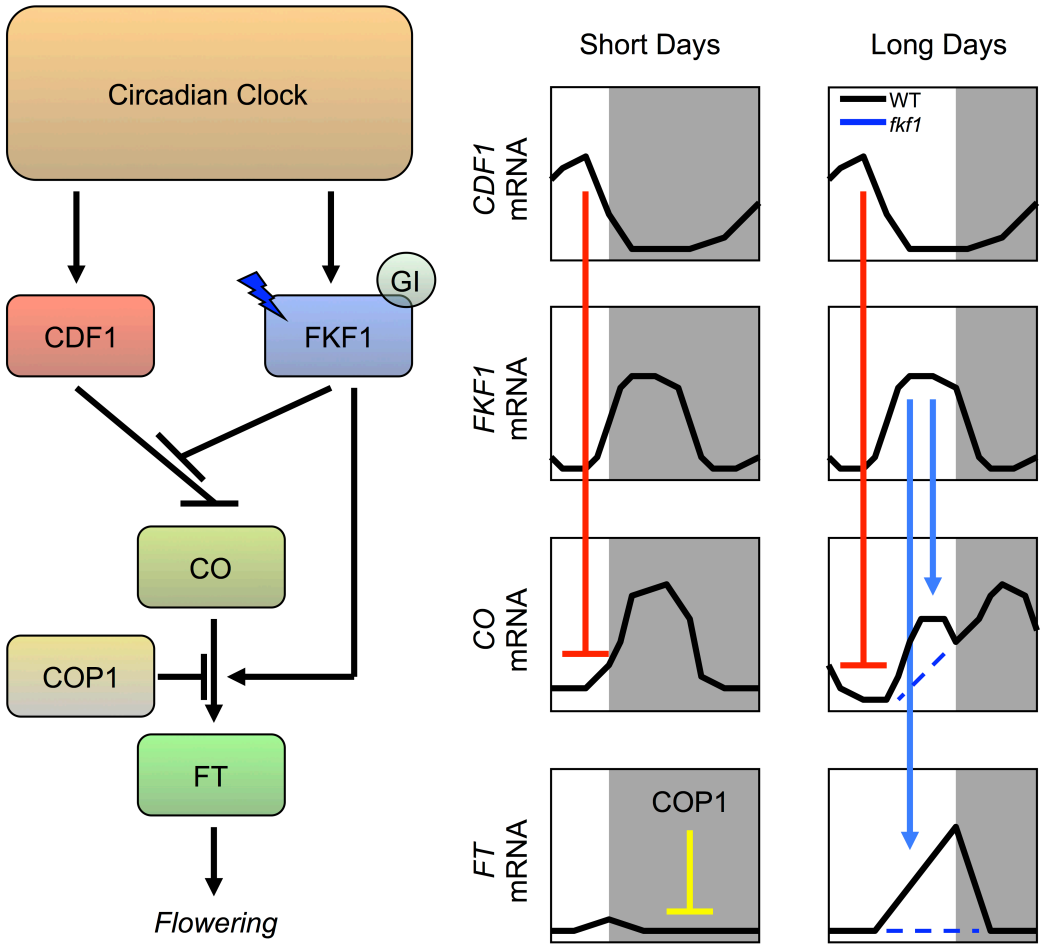


Figure 1.2: Schematic of photoperiodic flowering pathway. (Left) Rectangles represent components of the flowering system. As described in the text, the circadian clock regulates flowering through transcription of *FKF1* and *CDF1*. The FKF1 protein forms a blue light-dependent complex with GI to degrade CDF1 protein that inhibits transcription of *CO*. CO protein, in turn, promotes *FT* expression. However, due to strong degradation of CO protein at night by COP1, *FT* transcription is maintained at basal levels in SDs. The model derived hypothesis that GI-FKF1 forms a feed-forward network enhancing the CO-dependent activation of *FT* transcription is included [82]. (Right) Example of external coincidence. The timing of FKF1 rhythms coinciding with light in long days (LDs) promotes a peak of *CO* expression before dusk that is not seen in short days (SDs). In *fkf1* transgenic plants, this end-of-day expression of *CO* is lost. The increase in *CO* mRNA prior to dusk coincides with the large increase of *FT* levels that accelerate flowering in LD growth conditions.

all photoperiods (ZT8-10, ZT = zeitgeber time where ZT0 = dawn; [34, 80, 83]). The importance of this complex can be seen in *fkf1* loss-of-function mutations that leads to the loss of the dusk peak of *CO* mRNA and greatly reduced levels of *FT* mRNA

(Figure 1.2; [74]). The mechanism that regulates the second, night-time peak of *CO* mRNA is currently unknown, although CONSTITUTIVELY PHOTOMORPHOGENIC 1 (COP1), FLOWERING BHLH (FBH) and LONG VEGETATIVE PHASE 1 (LOV1) have all been implicated in *CO* transcription at this time [34, 84, 85]. This mechanism will be discussed in Chapter 4.

The photoperiod-dependent regulation of *FT* expression occurs due to the strong regulation of CO protein at night [86, 87]. During the day, CO protein follows a rhythm that resembles the rhythm of *CO* mRNA [86]. Importantly for the regulation of *FT* transcription in diurnal cycles, CO protein is strongly regulated by COP1 at night that targets CO for degradation [87, 88]. Thus, in SDs, the strong degradation of CO protein prevents accumulation of *FT* mRNA. However, in *cop1* loss-of-function mutations, *FT* mRNA is able to be expressed at night in SDs following a similar rhythm to *CO* mRNA, leading to an acceleration of flowering [34, 87]. Taking all these observations together, the external coincidence mechanism that regulates photoperiod-dependent *FT* expression occurs in two steps: 1. circadian regulated GI and FKF1 protein levels coincide with blue light signals in LDs that lead to their interaction and the activation of *CO* transcription; 2. CO protein accumulates during the day but is degraded at night by COP1 leading to *FT* mRNA peaking at dusk. In SDs, GI and FKF1 protein levels do not accumulate early enough in the light period to trigger activation of *CO* and *FT* mRNA (Figure 1.2).

A simple mathematical model of this system was constructed to determine whether there were any hidden components that are needed to be taken into account [82]. This is similar to what has been discussed previously in the example of the circadian clock models. The photoperiodic flowering pathway was added to the 'Locke2005' clock model such that *CO* mRNA was regulated by FKF1 in a light-dependent manner and a second, clock-dependent mechanism that maintains transcription at night in LDs. By including the strong dark-dependent degradation of CO protein, this model was able to describe the photoperiod-dependent differences in *FT* mRNA rhythms of WT plants [82]. However, in trying to simulate the *fkf1* loss-of-function mutation, the model simulated higher levels of *FT* mRNA than had been observed experimentally

despite recapitulating the loss of the first LD peak of *CO* mRNA [74, 81]. Thus, the hypothesis derived by the model to explain why simulations could not describe the data in the *fkf1* background was that FKF1 protein must play a second role in the system, namely in aiding CO to activate *FT* transcription (Figure 1.2). Thus, FKF1 protein plays two roles in the model, forming a coherent type 1 feed-forward network by positively regulating both *CO* and *FT* transcription resulting in an amplification of *FT* mRNA in LDs [82, 89].

However, since this model was constructed, a number of important results have been elucidated experimentally that add more complexity to the flowering system. First, the regulation of *CO* transcription by the GI-FKF1 complex has been shown to be indirect. The GI-FKF1 complex targets CYCLING DOF FACTOR (CDF) proteins for degradation that, in turn, inhibit *CO* transcription (Figure 1.2; [83, 90]). Second, GI has been shown to be active on the *FT* promoter to activate transcription [91]. Thus, similar to the proposed feed-forward network of FKF1 above, GI plays two roles in the flowering system activating *CO* and *FT* mRNA (Figure 1.2). Finally, transcription of *FKF1* and *CDF* components in the flowering system has been shown to be regulated by the circadian clock (Figure 1.2; [50, 78]). Thus, the circadian clock regulates *CO* mRNA and the flowering pathway through three pathways: *GI*, *FKF1* and *CDF* transcription. These new connections and components of the system will be the topic of Chapters 3 & 4.

In the discussion above, I have concentrated on the circadian regulation of plant physiology that enables plants to sense changes in day-length. However, light- and temperature-signalling also play important roles in determining plant physiology – the regulation of the circadian clock and output pathways by these environmental cues will make up the discussion in the rest of this chapter, introducing research of Chapters 4 & 5.

Regulation of circadian-controlled plant physiology by light: Photoreceptors

Arabidopsis thaliana has multiple photoreceptors that are able to respond to light of varying qualities (or wavelengths). These include the phototropins that respond to UV-A/blue light, cryptochromes and the ZTL protein family that respond to blue light, and phytochromes that respond to red & far-red light. Recently, a further UV-B photoreceptor has been found, UV RESISTANCE LOCUS 8 (UVR8) [92]. For a review of photoreceptors and their control of plant development, please see [93]. Here I shall concentrate the discussion to the roles of the ZTL protein family and the phytochromes in regulating the circadian clock and flowering.

Blue light sensors: the ZTL protein family

Within the ZTL family are three proteins: LOV KELCH PROTEIN 2 (LKP2), ZTL/LKP1/ADAGIO 1 (ADO1) and FKF1. These proteins have recently been reviewed by [94]. The structure of these three proteins show similarities in that they posses a Light-Oxygen-Voltage (LOV)/PAS domain in their N-terminus, followed by an F-box and Kelch repeats in the C-terminus [95]. A schematic of the FKF1 protein structure is shown as a representative of this protein family in Figure 1.3. The LOV domain is shared with the phototropin family that respond to blue light. Like the phototropins, the LOV domain of FKF1 binds a flavin chromophore that absorbs blue light [74]. However, this view is complicated by the observations that the ZTL family has been shown to play roles in red light signalling [74, 95–98]. This could be the result of a direct interaction with the red light-sensing phytochrome B (phyB). phyB has been shown to interact with proteins containing PAS domains, however interactions with ZTL, FKF1 & LKP2 through their PAS-related LOV domains has not been proven (Figure 1.3; [97, 99, 100]). Through the F-box domain, ZTL, FKF1 & LKP2 have been shown to interact with Arabidopsis Skp1-like (ASK) proteins to form Skp/Cullin/F-box (SCF) E3 ligases that target proteins for degradation [101, 102]. Thus, the specific functions of the ZTL family proteins may be regulated by light-dependent interactions.

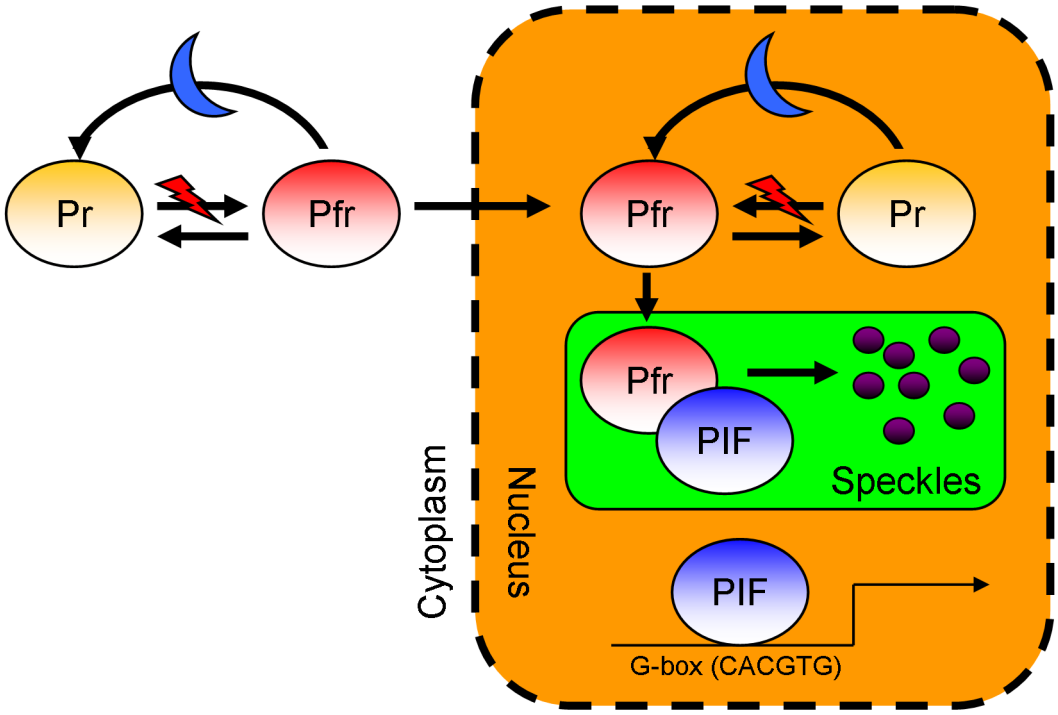


Figure 1.3: Representative protein structures of light regulated protein families. Schematics of protein structures for FKF1, phyB and PIF3 were obtained from the UniProt web service [103]. FKF1 contains a blue light-regulated LOV/PAS domain, an F-box that mediates interactions with SCF ligase complexes and kelch repeats (see text). phyB contains a red light-regulated PAS domain (see text) in between a GAF domain and a histidine-kinase (HK) domain. The GAF domain plays a role in the light-regulation of phyB signalling, whilst the HK domain is dispensable for phyB function [104, 105]. PIF3 contains a PAS and bHLH domain (see text). Note: UniProt does not show a PAS domain in the PIF protein structure; however, the PAS domain has been highlighted experimentally [99, 106].

As mentioned above, two members of this protein family play a role in the circadian clock and flowering systems. In the circadian clock, degradation of TOC1 protein has been shown to occur in the dark in a ZTL-dependent manner through a direct ZTL-TOC1 interaction [32, 102]. Similarly, PRR5 protein levels show the characteristics of dark-dependent degradation as PRR5 protein is stable in dark-grown *ztl* mutants [107]. Conversely, ZTL has been shown to stabilise GI protein by forming a blue light-dependent complex through the LOV domain of ZTL that alters the cellular distribution of GI, protecting GI from ELF3- and COP1-mediated destabilisation [33, 34, 108]. Thus, circadian rhythms in *ztl* loss-of-function transgenic plants have a long period phenotype [32, 95, 109]. Interestingly, period phenotypes have been observed when perturbing levels of the other ZTL family

members. Overexpression of LKP2 leads to arrhythmia in constant conditions and, similarly to ZTL, this could be due to an interaction with TOC1 protein [98, 102]. Further, FKF1 protein has been found to interact with both PRR5 and TOC1 leading to minor period lengthening phenotypes in *fkf1* loss-of-function mutants [109]. Despite this interaction, though, ZTL has been found to be the key regulator of PRR5 protein degradation with FKF1 and LKP2 playing minor roles in comparison, possibly as ZTL is more highly expressed [109]. Period phenotypes of *fkf1* and *ztl* loss-of-function mutants seem to be additive or synergistic, whilst loss of LKP2 function in *ztl;fkf1* double mutants does not alter period length, but dampens the amplitude of rhythms in constant light [109]. These phenotypes suggest the existence of a complex network between these family members that have been shown to interact with each other [110]. By altering the period lengths of circadian rhythms in constant conditions, the ZTL proteins are important components in the circadian clock to maintain 24hr rhythms, playing a light-dependent function in the regulation of TOC1, PRR5 and GI proteins.

In photoperiodic flowering, though, the formation of blue light-dependent complexes has proven to be crucial for the correct timing of flowering. By forming a complex with GI, FKF1 protein is stabilised in LDs and targets CDF1 protein for degradation in the nucleus [83, 90, 110]. Thus *fkf1* loss-of-function plants have delayed flowering phenotypes in LDs [74]. Additionally, rhythms of *CO* mRNA in *fkf1;ztl;lkp2* triple loss-of-function mutants mimic the low levels seen when GI function is absent supporting the view that the ZTL protein family regulates flowering through GI [83]. Interestingly, in contrast to the late-flowering *fkf1* mutants, *ztl;lkp2* double loss-of-function transgenic plants have early flowering phenotypes that are specific to SDs and not LDs [110]. Furthermore, this phenotype is dependent on levels of FKF1 in transgenic lines, suggesting that FKF1 acts downstream of ZTL and LKP2 to coordinate flowering. Intriguingly, similarly to GI protein, ZTL and LKP2 have been implicated in regulating FKF1 protein levels, such that increased FKF1 levels in SD grown *ztl;lkp2* transgenic lines provide the early flowering phenotype that is observed [83, 110]. Given the experimental observations outlined above, regulation of flowering by ZTL and LKP2 seems to occur in two stages. First, circadian

regulation of *FKF1* transcription (the only member of the family with rhythmic mRNA levels) is altered by the removal of ZTL and LKP2 that play roles in the circadian clock. Second, FKF1 protein stability is regulated by ZTL and LKP2, in conjunction with GI. As mRNA time-series data for the key flowering genes, *CDF1*, *FKF1*, *CO* and *FT*, as well as FKF1 protein profiles, have already been made available in different photoperiods and in different genetic backgrounds (such as WT, *fkf1*, *ztl*, *lkp2*, *ztl;lkp2* and *fkf1;ztl;lkp2*) it may be possible to test the hypothesis that ZTL/LKP2 regulate flowering through feed-forward regulation of FKF1 by extending the models presented in Chapters 3 & 4 [83, 110]. Interestingly, and in support of the model-derived hypothesis above postulating an interaction between CO and FKF1, LKP2 has indeed been shown to interact with CO protein suggesting that the formation of blue light-dependent complexes may be crucial at multiple levels of the flowering pathway [111].

Red light sensors: the phytochromes

In Arabidopsis, there are five phytochromes (phys): phyA-E [112]. Each of these proteins has different kinetics in response to light and differentially regulate plant development (for reviews see [113, 114]). Here I shall concentrate on the key red light-sensing phytochrome, phyB. The phyB protein can take one of two forms, either inactive Pr or active Pfr. Upon absorption of red light by the Pr form, phyB conformationally changes into the Pfr state (Figure 1.4; [115, 116]). Once in the Pfr conformation, phyB is able to enter the nucleus to form speckles that regulate downstream responses (Figure 1.4; [117–119]). Studies have shown that the nuclear speckle formation is light dependent and that the light-absorbing chromophore is required for this response. When the plant is then transferred into darkness, phyB presence in the nucleus decreases and the active Pfr form of phyB reverts back to inactive Pr through a process known as dark reversion [116, 120].

A key family of proteins required for phytochrome signalling are the bHLH PIFs (a schematic of the PIF3 protein structure is shown in Figure 1.3 as a representative of this protein family). Notably, transcription of *PIF4* and *PIF5* is regulated by the

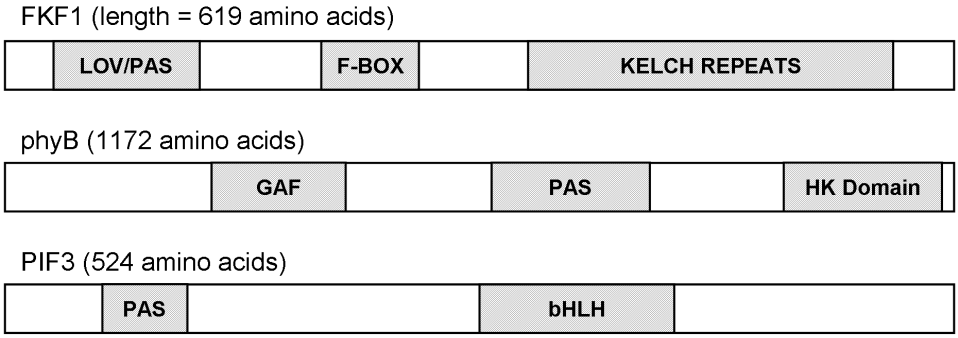


Figure 1.4: Schematic of the phyB-PIF network. Upon illumination by red light (red bolts), the inactive Pr form of phyB changes conformation to the active Pfr state. This reaction is reversible by growing plants in far-red light and through dark-reversion (blue moons). Once in the Pfr form, phyB is able to enter the nucleus and form nuclear speckles. By forming speckles, phyB is able to function through interacting with PIFs leading to their mutual degradation. The interaction between phyB and PIFs prevents the PIFs from binding to G-box motifs in target promoters and activating transcription.

circadian clock (see Chapter 2) and the protein levels are controlled by light to redundantly initiate hypocotyl elongation in SDs [61]. Thus, regulation of hypocotyl elongation by PIF4 and PIF5 forms a similar external coincidence system to photoperiodic flowering through a CO-dependent pathway presented above. Further to this system, PIF4 and PIF5 have been implicated in hormone- and sucrose-signalling, PIF5 has been implicated in the circadian clock through TIC (see above), PIF4 is involved in warm temperature-regulated plant development (~27°C, see below), PIF4 and PIF7 have been found to regulate the *CBF*-dependent cold acclimation pathway, while PIF1 and PIF3 have been shown to control Reactive Oxygen Species (ROS) promoting plant greening and regulating cell death [43, 121–130]. PIF proteins are able to regulate this diverse range of molecular mechanisms by binding to G-box (CACGTG) motifs in target promoters as either homo- or hetero-dimers (Figure 1.4; [131–133]). Thus, through the regulation of PIF protein levels, light signalling is able to interact with a range of other molecular pathways that control plant development.

However, the mechanism by which red light, and notably phyB, regulates PIF activity is complex. phyB contains PAS domains within the protein structure that are

crucial for the proteins function (Figure 1.3). Importantly, removal of the PAS domains prevents phyB from forming nuclear speckles required for responses to red light [134]. Thus, the conformational change between Pr and Pfr is believed to alter the accessibility of the phyB PAS domains to signalling partners. The phyB-PIF interaction occurs reversibly with changes in light conditions, such that the active Pfr form of phyB preferentially interacts with PIF1, 3, 4, 5 and 7 (Figure 1.4; [135, 136]). Furthermore, this interaction requires the PAS domain of PIFs to take place (Figure 1.3; [99]). This interaction is particularly important when de-etiolation takes place and plants are exposed to light for the first time. Under these conditions PIF-mediated transcriptional activation occurs acutely (i.e. within a few hours) due to the fast degradation of PIFs through their interaction with phyB (Figure 1.4; [132, 137]). Interestingly, phosphorylation of PIFs does not affect their interaction with phyB, but does regulate the degradation kinetics of both PIFs and phyB [138]. However, over longer durations of red light (i.e. for 4 days of continuous light), the PIFs are able to promote the degradation of phyB through a COP1-dependent mechanism (Figure 1.4; [137, 139]). Furthermore, over diurnal cycles protein levels of phyB remain relatively stable suggesting that the changes in phyB conformation are the key regulators of red light signalling under these conditions [140]. To further complicate this interaction, a recent study has highlighted that phyB is able to prevent the binding of PIFs to target genes through sequestration (Figure 1.4; [141]). Thus, not only does phyB regulate PIF degradation in the control of light signalling, but it is also able to inhibit PIF-mediated transcriptional activation.

As phyB is able to regulate transcription, a number of genes have been found to have altered responses to red light upon de-etiolation in transgenic plants with *phyB* loss-of-function mutations [142]. Importantly for the circadian clock, gene expression of *CCA1*, *LHY*, *GI*, and *PRR9*, as well as the clock regulated *CHLOROPHYLL A/B BINDING PROTEIN 2* (*CAB2*), experiences an acute ‘burst’ of expression that is reduced in *phyB* mutants [23, 142–145]. Recent studies have proposed that the acute ‘bursts’ of transcription in response to red light occur due to histone modifications in the chromatin of these genes [145]. However, this is not the only effect of phyB on circadian regulated gene expression. Under diurnal cycles with red light, phyB

appears to regulate the amplitude of circadian regulated gene expression [143]. Furthermore, at fluence rates (light intensity) above ~1μmol m^{-2} s^{-1}, *phyB* loss-of-function transgenic plants have long period phenotypes of circadian rhythms under constant red light [9, 10, 146]. The molecular mechanism used by phyB to regulate the circadian clock has not yet been elucidated – this will be the focus of Chapter 5.

Although phyB has been shown to regulate the circadian clock, studies have also found that phyB regulates flowering and that this process occurs downstream of the circadian clock. Across all photoperiods (SDs, LDs and constant light) flowering is accelerated in *phyB* loss-of-function plants [146–150]. The early flowering phenotype occurs in the presence of red light, such that phyB does not play a role in the blue light flowering pathway [147, 148]. A recent study has analysed roles of phyB and the circadian clock components CCA1 & LHY using loss-of-function transgenic plants [150]. In doing this, the results showed that the effects of removing *phyB* are able to suppress the delayed flowering phenotype of *cca1;lhy* transgenic plants grown in constant light. This supports a model where phyB acts downstream of the circadian clock. This mechanism fits in with studies that have shown that CO protein is strongly light regulated [86, 151–153]. In LD conditions, CO protein is expressed at a higher level in *phyB* loss-of-function transgenic plants compared to WT. Similarly, when plants are grown in far-red or blue light, where phyB signalling is minimal, the CO protein is stabilised to increase *FT* expression [151, 153]. Thus, light regulation of CO protein plays a role in determining the time taken for plants to flower.

Regulation of the circadian and flowering systems by temperature

Molecular pathways regulating temperature signalling

Similar to light signalling that regulates plant growth over a wide range of wavelengths through a variety of different molecular pathways, temperature is able to regulate plant growth and development by the use of multiple mechanisms (for a review see [154]). For example, in cool and freezing temperatures, a molecular cold

acclimation pathway is activated that helps Arabidopsis survive in the cool. Critical to this pathway is the expression of *CBF*s that positively regulate transcription of the *COLD REGULATED* (*COR*) family of genes [155, 156]. In response to cool temperatures, the *CBF* genes are activated by INDUCER OF CBF EXPRESSION 1 (ICE1) that is targeted for degradation by HIGH EXPRESSION OF OSMOTICALLY RESPONSIVE 1 (HOS1) [157–159]. In warmer ambient temperatures (~27°C), one of the key signalling components is PIF4 [122–125, 129]. Both transcription and protein levels of PIF4 increase with warmer temperatures and, furthermore, in *pif4* loss-of-function plants, flowering and hypocotyl elongation phenotypes in SDs are insensitive to increases in temperature. Thus, changing the temperature leads to the activation of different signalling mechanisms.

Interestingly, the circadian clock has been shown to be a key regulator of both of these pathways [20, 61, 64, 160]. This allows for the gating of temperature responses such that these pathways are more sensitive to temperature changes at specific times of day [156, 161]. From the gating experiments, it seems that expression of *PIF4* mRNA is more sensitive to warm temperatures at night, whilst *CBF* expression is more sensitive to cool temperatures during the day [156, 161]. A further advantage of these pathways being regulated by the circadian clock is that transcription of *CBF*s and *PIF4* are sensitive to changes in photoperiod such that both are expressed to a higher level in SDs compared to LDs at the times when they are sensitive to temperature, i.e. *CBF* expression is higher during the day, while *PIF4* expression is higher at night [128, 161, 162]. This leads to similar plant phenotypes, such as freezing tolerance and hypocotyl elongation, if plants are grown in either SDs or in temperatures that promote *CBF* and *PIF4* expression. The circadian clock has been proposed to regulate transcription of *CBF*s and *PIF4* through CCA1 & LHY, the PRRs and the EC [25, 31, 64, 160]. Recently, the use of mathematical model selection techniques (presented in Chapter 2) has elucidated that, as well as CCA1 & LHY, TOC1 is able to directly regulate *CBF3* expression [163]. The focus of Chapter 2 is to investigate the mechanism used by the circadian clock to regulate *PIF4* and *PIF5* expression. The resulting model is used as part of a larger system in Chapter 4.

Temperature regulation of the circadian clock

Whilst temperature regulation of plant physiology downstream of the circadian clock is relatively well understood, how the circadian clock mechanism responds to varying temperatures whilst maintaining temperature compensation (see above) is not known. In constant light, temperature is able to entrain the circadian clock to H:C cycles [7]. Furthermore, a number of transgenic lines have been shown to have altered temperature compensation phenotypes. Salome *et al.* found that *prr9;7* double loss-of-function transgenic plants have a period that increased linearly with temperature [7]. This phenotype can be suppressed by decreased levels of *CCA1* and *LHY* mRNA suggesting that the PRR regulation of *CCA1* & *LHY* transcription is important for temperature compensation. This view is supported by the observation the *PRR9* and *PRR7* are acutely induced by heat shocks (37°C) given in constant dark conditions independently of the presence of *CCA1* or *LHY*. The acute activation of *PRR9* & *7* is then able to trigger circadian rhythms of *CCA1* & *LHY* in constant darkness [164]. An alternative theory of temperature compensation has recently been investigated. Analysis of *gi* and *lhy* loss-of-function transgenic plants has led to the 'network balancing' hypothesis whereby temperature induced period lengthening due to alterations in one area of the circadian clock need to be counter-balanced by period shortening effects elsewhere in the mechanism to maintain a 24hr period [6]. The 'network balancing' hypothesis was recently tested mathematically and showed that temperature compensation results in temperature-dependent increases of LHY protein levels when plants are grown in blue light [8]. Interestingly, the temperature-sensitive mathematical model of the circadian clock only used the Arrhenius equation, which describes the temperature-dependency of chemical reactions, to alter specific light-regulated parameters [8]. This suggests that temperature regulates the circadian clock through specific (light regulated) mechanisms, rather than through a global alteration of all reaction rates with temperature changes.

The finding that LHY protein levels change with temperature fits in with another avenue of research which shows that transcription of circadian clock components is under the control of temperature-dependent alternative splicing. The morning loop

genes *CCA1*, *LHY* and the *PRR*s have all been shown to have temperature-associated alternative splicing events [165]. The results of alternative splicing is that effective *LHY* mRNA levels decrease with temperature changes from 20°C to 4°C, whilst *CCA1* transcript obtains higher expression upon the transition from 20°C to 4°C, whilst plants acclimated to 20°C or 4°C have similar expression levels [165]. As a consequence, altering the spliceosome mechanism can lead to the loss of temperature compensation [166]. Recently, this result has been extended whereby the protein products of two *CCA1* splice variants have been shown to negatively regulate each others function in a temperature dependent manner [167]. Similar mechanisms have been observed in other species such as *Neurospora crassa*, the model fungus, and *Drosophila*, the model fly [168, 169]. In Chapter 5, potential mechanisms through which the circadian clock may be regulated in warm ambient temperatures (~27°C) will be discussed.

Temperature regulation of flowering

As with many plant phenotypes, flowering is sensitive to changes in temperature. In response to drops in temperature from ~23°C to 16°C or 4°C, flowering of LD-grown plants is delayed [170, 171]. Changes in CO protein stability are important for the observed delayed flowering phenotype. Studies have observed that, under LD conditions, CO protein is increased in *hos1* loss-of-function transgenic plants [170–172]. However, it is interesting to note that flowering phenotypes of the *hos1;co* double loss-of-function mutation may be partially additive, suggesting that HOS1 is able to regulate flowering through CO-independent mechanisms [171]. When temperatures are increased from 22°C to 27°C, flowering is accelerated when plants are grown in SDs [173, 174]. This phenotype has been observed to be dependent on the presence of PIF4, such that the loss of PIF4 function leads to a temperature insensitive flowering phenotype [174]. This mechanism is proposed to occur due to chromatin modifications that allow PIF4 to regulate transcription of *FT* specifically at 27°C compared to 22°C, when the presence of H2A.Z nucleosomes prevents PIF4 activity on the *FT* promoter. Interestingly, as with LD-grown *co* loss-of-function plants in LD conditions, the phenotype of SD grown *co* plants are also temperature

sensitive in the warm [173]. However, when plants are grown in the warm in LDs, the phenotypes of *co;PIF4-ox* are partially additive, suggesting that, as with HOS1, PIF4 may regulate flowering partially through a CO-dependent mechanism [174]. In Chapter 4, the regulation of *FT* transcription and flowering due to temperature increases through a PIF4-dependent mechanism shall be researched.

Interaction of light and temperature signals

As suggested by the role of PIFs in red light- and temperature-signalling, and the regulation of *PIF* transcription by the circadian clock, there is crossover between light- and temperature-mediated responses. Here, I shall discuss some of the published observations that relate phyB- and temperature-signalling, particularly in the regulation of flowering. Experimental observations have shown that flowering phenotypes of *phyB* loss-of-function transgenic plants are more sensitive to increases in temperature than phenotypes of WT plants grown in SDs [149]. This occurs through an increase in *FT* mRNA in warm-grown *phyB* transgenic plants, whilst other components of the flowering time system remain relatively constant across temperatures in both WT and *phyB* plants. This would support the role of light signalling and temperature in the regulation of a direct regulator of *FT* mRNA. Interestingly, the diurnal regulation of *PIF4* mRNA, an activator of *FT* transcription, has been shown to increase at night in LDs with temperature [161]. This night-time increase of *PIF4* mRNA is subsequently increased by the loss of phyB activity, supporting the theory that part of phyB function is to suppress responses to warm temperature. Additionally, an increase of *PIF4* and *PIF5* mRNA as a result of 37°C heat shocks given in constant dark is greatly decreased in *cca1;lhy* loss-of-function plants, supporting a role for the clock in the temperature regulation of *PIF* transcription [164]. An important difference between the development of transgenic loss-of-function *phyB* and WT plants is the ability of the plants to sense diurnal changes in light and temperature [146, 175]. In this scenario, WT plants are able to sense changes in day-time and night-time temperature, but *phyB* transgenic plants sense a constant temperature throughout a whole 24hr day. Furthermore, *phyB* loss-of-function mutants are less sensitive to changes in photoperiod [146]. Again, this

suggests that phyB plays an important role in a plants response to diurnal cycles. Similarly, studies of freezing tolerance have highlighted a role of phyB in suppressing freezing tolerance by decreasing *CBF* and *COR* expression through light-signalling components PIF4, PIF7 and LONG HYPOCOTYL 5 (HY5) [128, 176, 177]. Whilst these examples highlight a link between red light- and temperature-signals in the regulation of plant development downstream of the circadian clock, it is unknown how these signals regulate the circadian clock and whether these signals converge on a single mechanism. This will be discussed in Chapter 5, where data is presented showing how temperature and phyB regulates the circadian clock.

Thesis Summary

In this chapter I have introduced how the circadian clock and flowering of the model plant, *Arabidopsis thaliana*, is regulated by photoperiod-, red & blue light- and temperature-dependent signals from the external environment. Over the course of this thesis I will develop on the work that has been discussed in this chapter to investigate the interaction between photoperiod-, red & blue light- and temperature-dependent plant development, in particular flowering.

In Chapter 2, theoretical techniques will be introduced that can be used to elucidate how the circadian clock regulates WT expression rhythms of output genes. As an example the regulation of *PIF4* and *PIF5* transcription will be discussed, however the same technique has been used to determine the regulation of *CBF* gene expression in a collaborative ROBuST project [163]. From this analysis, I shall show that the simplest mechanism that describes transcriptional regulation accurately across multiple photoperiods requires more than one circadian clock component.

In Chapter 3, I will develop the mathematical model of flowering discussed above to incorporate the role of CDF1 that inhibits transcription of *CO* mRNA. Furthermore, new data obtained by collaborators is presented showing that CDF1 and FKF1 are able to directly regulate *FT* transcription, with FKF1 stabilising CO protein in a blue

light-dependent manner. The resulting mathematical model takes into account these forms of regulation and then analyses the role of FKF1 in flowering. This model predicts how *FT* mRNA is altered in *FKF1-ox* transgenic plants; determines which of the FKF1-CDF1 or FKF1-CO protein interactions is relatively more important for the correct regulation of *FT* mRNA, and; whether the circadian regulation of FKF1 is required to delay the accumulation of *FT* until dusk. This model will, therefore, highlight the important role of blue light-signalling components in flowering.

In Chapter 4, the flowering model presented in Chapter 3 will be extended further to include the circadian regulation of *CDF1* and *FKF1* mRNA, as well as the temperature regulation of *FT* mRNA by PIF4. By building a more complex model, hypotheses will be generated that link components of the circadian clock, namely ELF3, CCA1 and LHY, to important roles in flowering that are downstream of *CDF1* and *FKF1* transcription. This suggests that GI is not the only clock component to form protein complexes that act in output pathways of the circadian clock. Furthermore, by analysing how *FT* mRNA is regulated by PIF4, PIF5 and temperature in LDs, simulations will propose that CO protein is an important temperature sensor and the warm-induced acceleration of flowering can be modelled through an interaction between PIF4 and CO. Thus, this model incorporates components of both red light- and temperature-signalling in the regulation of flowering.

In Chapter 5, I will discuss the regulation of the circadian clock by phyB- and temperature-signalling. By manipulating the most recent published circadian clock model to match rhythms of circadian clock components obtained experimentally in *phyB* transgenic plants at 22°C, I will hypothesise two potential mechanisms through which phyB plays a role in the clock machinery. Further, by presenting data from WT and *phyB* loss-of-function plants at 27°C, I will highlight potential mechanisms that relate phyB and temperature signalling. This represents the first attempt to mathematically unravel the mechanisms within the Arabidopsis circadian clock that are able to sense changes in red light-dependent phyB signalling and temperature together.

Finally, in the last chapter, the presented results will be discussed to show how they further our understanding of the mechanisms regulating plant development in response to changes in the external environment.

Chapter 2: Model Selection Techniques & the Circadian Clock

The techniques discussed in this chapter are being published as part of:
J. Keily, *et al.* (2013) Model selection reveals control of cold signalling by evening-phased components of the plant circadian clock. *Plant J.*, vol. 76, pp. 247-57.

As discussed in the introduction chapter, mathematical models of the Arabidopsis circadian clock have evolved from a single negative feedback loop between CCA1 & LHY and TOC1 to larger models featuring three feedback loops and 10 components [12, 23, 36–38]. The consequence of including new components to the clock model is that the number of parameters that represent transcription/translation rates, degradation rates and protein binding affinities has increased from 29 to 104 [12, 23]. This increase in complexity could, theoretically, lead to the model 'overfitting' experimental training data. A model is said to 'overfit' data if a less complex model exists that is capable of matching the same data sets with a similar degree of accuracy. This has both technical and practical implications. When a model 'overfits' data it is hard to generate simple and experimentally testable hypotheses due to the extra complexity in the system, which could be characterised by a high number of hidden variables or complicated representations of biological mechanisms. It is therefore relevant, when modelling output pathways of the circadian clock – as we shall come to in the next chapters – to understand the properties of both simple and complex models, and whether this complexity helps improve the mathematical description of experimentally observed behaviours.

Circadian clock models have previously been extended to create larger external coincidence models that predict photoperiodic control of flowering (see Chapter 1; [82]). Another important clock regulated physiological response, and the focus of this study, is elongation of the hypocotyl, or seedling stem. Hypocotyl extension is regulated by PIF4 and PIF5 such that removal of both of these factors in a *pif4;5*

double loss-of-function transgenic plant eradicates the daily rhythm of hypocotyl elongation that peaks at the end of the night in wild type plants [61]. The mRNA of these transcription factors is regulated by the circadian clock leading to altered rhythms of *PIF4* and *PIF5* transcription in genetic perturbations of clock components, whilst the timing of maximum mRNA expression does not change across photoperiods [25, 61, 178, 179]. Studies have shown that *PIF4* and *PIF5* mRNA are directly regulated by the Evening Complex (EC) of the circadian clock [25, 179]. As stated in Chapter 1, the function of this complex is dependent on the availability of ELF3 to interact with LUX that binds to DNA at LBS motifs (GATWCG; [28]). Hence, knocking out ELF3 removes EC function. Furthermore, members of the PRR protein family (PRR7 & 5) have recently been shown to inhibit transcription of promoter regions containing G-box (CACGTG) motifs, including *PIF4* and *PIF5* [31]. However, the rhythmic expression of *PIF4* and *PIF5* transcription is not completely eradicated in *elf3* or *prr* mutants grown under diurnal conditions, showing similar phase shifts as other clock components [25, 27, 178, 179]. This suggests that the regulation of *PIF4* and *PIF5* mRNA by circadian clock components does not occur completely independently of each other.

In this chapter, by adding a *PIF4/5* component to two circadian clock models of differing complexity, simulations of *PIF4/5* mRNA will be analysed to determine whether added complexity to the circadian clock models improves their performance to describe *PIF4* and *PIF5* transcription across three photoperiods. The two previously published clock models used in this analysis are the 'Locke2005' two-loop model that contains 61 parameters and the 'Pokhilko2010' three-loop model that contains 90 parameters [36, 38]. The rationale for this choice is as follows: the 'Locke2005' model was the first clock model able to describe rhythms of *TOC1* transcription in the *cca1;lhy* mutant, and; the 'Pokhilko2010' model was the first model to accurately match details of circadian regulation observed across a range of photoperiods, such as the decrease of peak *CCA1* mRNA expression with increased photoperiod and double-peaked *CCA1* transcription in skeleton photoperiods [23, 36, 38, 80]. These two considerations are important for circadian clock models since these mathematical systems need to be able to accurately explain biological

observations from multiple conditions, such as genetic and/or environmental perturbations. Recently, a newer circadian clock model has been constructed that now includes the EC (the 'Pokhilko2011' clock; [12]) that is known to regulate *PIF4* and *PIF5* transcription [25]. However, it would be difficult to compare model selection techniques for two models where one includes the EC and the other does not as results may be skewed in favour of the EC-containing model. Similarly, as the 'Locke2005' model does not incorporate the PRR proteins, inhibition of *PIF4* and *PIF5* transcription by the PRRs shall not be directly modelled here [31]. In the models presented in this chapter I shall use TOC1, a PRR family member with similar timing of expression to components of the EC, which is present in both the 'Locke2005' and 'Pokhilko2010' clock models, as a representative evening suppressor of *PIF4/5* transcription. By comparing simulation accuracy of *PIF4/5* mRNA, I will be able to examine the advantages and disadvantages of model complexity. Furthermore, I will also aim to predict how *PIF4* and *PIF5* transcription is regulated by the circadian clock and whether regulation requires more than an evening repressor alone. In order to do this, I shall use two forms of model selection techniques that balance the accuracy of the model to experimental data against the complexity of the model (known as the frequentist approach) or the variability associated with a model that can result from the uncertainty in parameter values (the Bayesian approach). Both of these approaches are based on the calculation of likelihood probabilities - the higher the likelihood probability the more probable it is that this model is the best description of the data from the set of models tested [180–183].

The two methods that will be described here are the corrected Akaike Information Criterion (AICc) and an adapted version of Thermodynamic Integration (TI). The AICc is an unbiased estimator of the Kullback-Leibler divergence that determines the quantitative difference between experimental data and model simulations. AICc also penalises the accuracy of the simulations by the number of parameters that are in the model [180–182]. However, the use of the AICc as a model selection technique is constrained by the number of data points available to compare against model simulations such that there are more data points than model parameters (see Model

Selection Techniques in Methods). As such, to correctly analyse a complex model, a large number of data needs to be generated to use the AICc analysis. TI is generally used as part of parameter optimisation whilst creating models, sampling parameters from known probability distributions [183, 184]. Therefore, the model that is favoured by TI is the system and parameter set that generates simulations with the highest accuracy when compared to biological observations. Whilst this technique does not contain a term that directly penalises a model by the number of parameters in the system, TI is able to recognise how complex a model is by the number of dimensions in parameter space that require exploring to find an optimal parameter set for the system under question. Consequently, the resulting model probabilities from the analysis contain information about the complexity of the models tested.

In this chapter, I shall generate a range of possible models of *PIF4* and *PIF5* transcription as an output of the circadian clock. Then, by using model selection techniques, it will be possible to determine two things. First, whether WT *PIF4* and *PIF5* mRNA rhythms can be accurately described with simple clock models. Second, whether morning expressed clock components are required for *PIF4* & *5* transcription in conjunction with inhibition in the evening.

Methods

Promoter Analysis

As *PIF4/5* mRNA levels are subject to circadian regulation, the *cis*-regulatory database ATCOECIS was used to identify regulatory motifs located in the promoter regions of these genes [185]. This database combines information from two databases that are updated manually as new literature on *cis*-regulatory elements is published – PLACE and AGRIS [186, 187]. Circadian regulated promoter regions that were found through the database were checked manually in a region up to 3000 basepairs (bp) from the transcription start site of the gene or to the nearest gene. The promoter binding site search included the following: Morning Element (ME; AACCAC; [19]), Evening Element (EE; AATATCT; [19, 188]), CCA1 binding site

(CBS; AAATCT; [19]), LUX binding site (LBS; GATWCG; [28]) and the Hormone Up At Dawn sequence (HUD; CACATG; [189, 190]). Knowledge of the ME is limited, but the sequence is found to be overrepresented in genes with a phase (peak expression) near dawn and is particularly important when the EE is mutated [19]. The EE and CBS have both been shown to be bound by CCA1. When bound to EE, CCA1 has been proposed to act as a repressor, delaying transcription, such that these genes peak in the evening phase. Conversely, when bound to CBS, CCA1 appears to function as an activator, which leads to peaks of target gene expression at around dawn [19, 188]. Furthermore, EE-like motifs are over-represented in promoters targeted by TOC1 for inhibition [30]. The HUD sequence is conserved in the promoter regions of genes regulated by light and hormone signalling, known as phytohormones [189, 190].

Data Sources

Data were collected from various sources. Transcription profiles for *CCA1*, *TOC1*, *PIF4* and *PIF5* mRNA were obtained from the DIURNAL websource (http://diurnal.cgrb.oregonstate.edu/; [191]) that used high-density oligonucleotide array assays of *Arabidopsis thaliana* accession Landsberg *erecta* (L*er*) plants sampled under SD (8L:16D) and LD (16L:8D) conditions. Transcript profiles for *PIF4* and *PIF5* under 12L:12D were acquired experimentally, as described in Experimental Methods below (K. Sidaway-Lee, University of Exeter, unpublished). Transcript profiles for *CCA1* and *TOC1* under 12L:12D were obtained from [80]. All *PIF* data were normalized such that the peak of expression in each photoperiod was 1.

Experimental Methods

Experimental work was conducted by Dr. Kate Sidaway-Lee and Dr. Dana MacGregor, collaborators from the ROBuST project (Dr. Steve Penfield's group, University of Exeter, UK).

- *Gene Expression Analysis*

Seeds of *Arabidopsis thaliana* accession Landsberg *erecta* (L*er*, available from NASC, Loughborough, England) were sterilised, then germinated and grown in 12L:12D cycles at 25°C under 80 µmol m^{-2} s^{-1} white light on full MS agar (Duchefa, Ipswich, England), without added sucrose. Seedlings were harvested every 4 hours. The RNA was purified using the RNeasy Plant RNA isolation kit (Qiagen, Crawley, England) according to the manufacturer's protocol. First-strand cDNA was synthesized with 5 µg of total RNA in 20 µl reactions, Superscript II Reverse Transcriptase (Invitrogen, Paisley, Scotland) and oligo dT according to the manufacturer's instructions, and 180 µl water was added before the PCR step. Real-Time RT-PCR was performed with SYBR-green detection in an ABI7300 instrument (Applied Biosystems, Paisley, Scotland). 2 µl of the diluted cDNA template was used, along with the primers. Primers used were *PIF4* (AT2G43010: 5' – GTTGTTGACTTTGCTGTCCCGC – 3'; 3' – CGACTCAGCCGATGGAGATGTT – 5') and *PIF5* (AT3G59060: 5' – CGCCGGAGATCCAAATCCCAACAT – 3'; 3' – GCGGGAAATCAGACCGTGTGCAACAA – 5') normalized against the control gene UBQ10 (AT4G05320: 5' – CACACTCCACTTGGTCTTGCGT – 3'; 3' – TGGTCTTTCCGGTGAGAGAGTC – 5') [192].

- *Chromatin Immunoprecipitation (ChIP) assays*

Chromatin immunoprecipitation was performed following the protocol in [193] with modifications using plants expressing the pCCA1:CCA1-YFP protein construct that has been described previously [194]. Seedlings from the Col-0 (Columbia) accession were grown on ½ MS agar plates at 22°C for 14 days with 12L:12D cycles and harvested at ZT2. The chromatin was sheared to between 100 and 1,000 bp in a Bioruptor UCD 200 (Diagenode, Liege, Belgium) at high intensity for 10 minutes (cycles of 30s on / 30 s off) at 4°C after [195]. An aliquot of the chromatin was reserved at this point as the Input chromatin. Immunoprecipation used equilibrated Dynabeads Protein A (Invitrogen/Life Technologies, Paisley, UK). The pre-cleared chromatin was transferred away from the beads and incubated with rotation over

night at 4°C with a 1:1000 dilution of anti-GFP (Abcam ab290; Abcam, Cambridge, UK). A new aliquot of equilibrated beads was then added and incubated with the chromatin solution for 2 hours at 4°C with rotation and then washed with low salt, high salt, and lithium chloride washes. The immunocomplexes were recovered from the beads by boiling for 10 minutes in the presence of 10% Chelex resin (BioRad, Hemel Hempstead, UK) and the proteins removed using Proteinase K Solution (Invitrogen/Life Technologies, Paisley, UK) at 50°C. The reserved Input chromatin was also processed in parallel with Chelex and Proteinase K and then purified using QIAquick PCR purification Kit (Qiagen, Manchester, UK). qPCR on the ChIP and Input DNA was performed in triplicate using Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent, Wokingham, UK) on a Mx3005P machine. The results were calculated so that percent input was equal to $100*(primer\_efficiency^{dcT})$ where dcT is the difference between the adjusted input cT and the ChIP sample cT. The input cT was adjusted to account for the dilution factor of the input chromatin. The primer efficiency was unique to each primer pair and is equal to $10^{\frac{-1}{slope(cT=m*\log(Input\_concentration+b)}}$. Primers used in this experiment were *ACTIN7* (*ACT7*, AT5G09810: 5’ – GTATCGGGTGACAATGCAGCTATTA – 3’ and 3’ – TGCTGGAGTAAAACATAAGCCACTC – 5’), *PIF4* (AT2G43010: *PIF4-a*, 5’ – CCAATCTGCCGACAAGTTTC – 3’ and 3’ – ACACCGTAACACCATCACGA – 5’; *PIF4-b*, 5’ – CCACGTGTCGTTCATTTCAA – 3’ and 3’ – GATAGAGAGTTGTGTTGGGCG – 5’; *PIF4-N*, 5’ – CGGAGTTCAACCTCAGCAGT – 3’ and 3’ – CAATTCAGAACAATCCCGGT – 5’), and *PIF5* (AT3G59060: *PIF5-a*, 5’ – TAGGCCCAATAACGCATCTC – 3’ and 3’ – TATCGGTTTAGAAGATGATGGAA – 5’; *PIF5-b*, 5’ – GTCCCTCCTTGCTCGATTTT – 3’ and 3’ – TGGAGAGGGTTGTTTGGTTT – 5’; *PIF5-c*, 5’ – TGGAGAGGGTTGTTTGGTTT – 3’ and 3’ – TGACATGGAACAAGTGTTTGC – 5’; *PIF5-N*, 5’ – GGACATGTTGGGATTTGGAT – 3’ and 3’ – GACCACCGACAGTCTTCATATG – 5’). The *GI* primers, used as a positive control, are the same as those in [179].

Model Construction

Due to the similarities in the rhythms of *PIF4* and *PIF5* mRNA levels, and the phenotypes of their respective single mutants, the *PIF4* and *PIF5* components were grouped together [61, 196]. This simplification meant that only one equation needed to be added to each of the clock models to describe the circadian regulation of *PIF4/5*:

$$\frac{dc_{P45}^{(m)}}{dt} = n\prod_{I}\left(\frac{g_I^a}{g_I^a + c_I^a}\right)\prod_{A}\left(\frac{c_A^b}{g_A^b + c_A^b}\right) - mc_{P45}^{(m)} \tag{2.1}$$

where $c_{P45}^{(m)}$ represents the normalized level of *PIF4/5* mRNA, n is the transcription rate, I are inhibitors of *PIF4/5*, A are activators, a and b are Hill coefficients, g_I and g_A are constants of inhibition or activation and m is the degradation rate. As the EC has not been modelled in the 'Locke2005' or 'Pokhilko2010' models used in this study, we used the simulated *TOC1* mRNA rhythm as an approximation of EC activity due to its similar timing of peak activity at approx. ZT14 in a 24 hour cycle [25, 38, 80]. The choice of multiplicative regulation was due to the observation that the whole rhythm of *PIF4* and *PIF5* was not eradicated in an *elf3* mutant grown in diurnal cycles, suggesting that the components generating the rhythmic *PIF* gene expression are unlikely to act independently.

The models were optimised by comparing *PIF4/5* mRNA data from 12L:12D diurnal cycles (K. Sidaway-Lee, University of Exeter, unpublished) to simulated rhythms using a parallel genetic algorithm (PGA) in the model optimisation framework, SBSI Visual (http://www.sbsi.ed.ac.uk/; see Table 2.I). Model equations were solved and simulated using the differential equation solver CVODES [197, 198]. Parallel genetic algorithms converge to find a parameter set where the error between the simulation and data reaches a minimum [199]. In our optimisation process, we set the target minimum error to be 0.01. This minimum, though, could be a local minimum in the parameter space and not the global minima. Studies have discussed that using'sloppy' parameters, which are not necessarily the global minima of parameter

Table 2.I: Parameter values of models obtained from SBSI Visual. T = TOC1, C = CCA1. d = down-regulated/inhibited transcription, u = up-regulated/activated transcription. A:B implies multiplicative regulation between A and B.

	'Locke2005'				**'Pokhilko2010'**			
k	**Td**	**Td:Xd**	**Cu**	**Td:Cu**	**Td**	**Td:Xd**	**Cu**	**Td:Cu**
n	9.416	9.6	0.292	0.5	10	10	0.717	10
g_1	0.461	0.461	0.05	0.019	0.236	0.231	0.191	0.302
a	2	2	2	2	2	2	2	2
g_2	-	2.105	-	1.034	-	10	-	0.1607
b	-	2	-	2	-	2	-	2
m	10	10	0.281	0.564	8.236	8.021	0.766	6.267

space, can lead to models displaying relatively accurate dynamics of a system across a number of conditions [200]. Simulated annealing methods attempt to find global minima but at a greater computational cost compared to parallel genetic algorithms.

Studies have shown annealing procedures are only slightly better at finding global minima compared to genetic algorithms [201]. By optimising the models in SBSI Visual, parameter values can be estimated within a specific range. Models were constrained such that all parameters could take values between 0 – 10 nM or nM/hr [12, 23, 36, 38]. Hill coefficients were set to 2 as a result of dimer formation by circadian clock proteins. The initial condition for *PIF4/5* mRNA in the optimisation was taken to be the average initial value from the available data sets. Once optimised, the parameter sets that provided models with the smallest cost were used for further analysis.

Cost Function and Pearson Correlation

In the following sections we will use the following general notation: in $j = 1,...,D$ datasets there will be $i=1,...,m$ datapoints, $n_i^{(j)}$, evaluated at certain timepoints t_i; M is a model simulation from a model of k parameters that takes the values $M_i = M(t_i)$. The cost function is defined to be:

$$Cost = \sum_{j=1}^{D} \sum_{i=1}^{m} \frac{(n_i^{(j)} - M_i)^2}{(m_{(j)} + 1)} \tag{2.2}$$

where $m_{(j)}$ is the number of datapoints in dataset j. The lower the cost value, the more accurately the model simulations match the data.

The Pearson Correlation coefficient is defined as:

$$r_{DM} = \frac{1}{D} \sum_{j=1}^{D} \frac{\sum_{i=1}^{m} (n_i^{(j)} - \overline{n_i^{(j)}})(M_i - \overline{M_i})}{\sqrt{\sum_{i=1}^{m} (n_i^{(j)} - \overline{n_i^{(j)}})^2 \sum_{i=1}^{m} (M_i - \overline{M_i})^2}} \tag{2.3}$$

where $r_{DM} \in [-1,1]$ [202]. For a model to have a good rhythmic match with the data then $r_{DM} \approx 1$. A value of -1 would imply that the model has the opposite rhythm to the data.

Model Selection Techniques

Here I shall derive the AICc and TI techniques used in this study. As stated previously, TI is used to help calculate parameter values. However, since the circadian clock models have parameter values already determined, I have adapted the TI method to assume that all parameter values were obtained from a Normal distribution. As the AICc analysis is frequentist in its approach, and TI is a Bayesian method of analysis, I then create hybrid versions of AICc (that takes into account parameter uncertainty, $AICc^U$) and TI (that includes an explicit penalty term for a models complexity, TI^p) to determine if the results change.

- *Calculation of AICc*

The AIC is calculated as [180]:

$$AIC = -2\log(L_{MLE}) + 2k, \tag{2.4}$$

where L_{MLE} is the maximum likelihood estimate of the likelihood function. The model that provides the smallest AIC is then deemed to be the most likely model from a set of models. Hurvich & Tsai showed that the AIC, though, was not an accurate model selection measure when the ratio of datapoints to parameters was small [181, 182]. The AICc overcame this issue:

$$AICc = -2\log(L_{MLE}) + \frac{2q(k+1)}{q-k-2}, \tag{2.5}$$

where $q>k+2$ is the total number of datapoints used in the analysis, ie. $q = Dm$ [181]. As with the AIC, the smallest AICc value corresponds to the model that is believed to be most likely. Hence, the condition $q>k+2$ ensures that the second term of (2.5) is strictly positive and for larger values of k (or, smaller ratios of q:k) will increase the resulting value of AICc more so than for a smaller model with the same accuracy.

In (2.4) and (2.5) I have shown the formulations of AIC and AICc featuring the likelihood probability. As pointed out by Burnham & Anderson, these results can be reduced in the special case that we assume the differences between a datapoint from dataset j and the model simulation at the same time follow a normal distribution [182]. Hence,

$$n_i^{(j)} - M_i \sim N(0,\sigma^2), \tag{2.6}$$

where $\sigma^2 = \frac{\sum_{i=1}^{m}(n_i^{(j)} - M_i)^2}{m}$. Using this assumption, the first term of the AIC and AICc reduces to

$$-2\log(L_{MLE}) = m\log(\sigma^2). \tag{2.7}$$

Thus, the reduced AICc is

$$AICc = \sum_{j=1}^{D} m_{(j)} \log(\sigma_{(j)}^2) + \frac{2q(k+1)}{q-k-2}. \tag{2.8}$$

For this analysis to work, though, it is crucial for $q>k+2$. This means that we need a large number of data for this analysis since k can be as large as 96 for the largest *PIF4/5* model. However, genes with circadian regulation should have the same level of expression at the same time of day when entrained in diurnal conditions. This means that circadian data series generated from qPCR can be repeated to cover a greater period of time (the same applies for model simulations of an entrained limit cycle system). Therefore, $n^{(j)}(t_i) = n^{(j)}(t_i + 24(d-1))$, where d is the number of whole days. Hence, we can take a dataset that describes 1 day to describe 3 days by repeating the dataset 3 times. For example, if we have data at *t=0, 4, 8, 12, 16, 20, 24* from 1 day with *n(0)=n(24)* then we count the points in *t=[0,20]* three times to create three limit cycles and the *t=24* once as the final timepoint of the 3 days. Since the model simulates circadian gene expression on a limit cycle, the simulations at the respective timepoints will also be the same in day 1 as day 3.

Thus the final AICc used in this study is:

$$AICc = d\sum_{j=1}^{D}(m_{(j)}-1)\log(\sigma_{(j)}^2|_{t\neq 24}) + \sum_{j=1}^{D}\log(\sigma_{(j)}^2|_{t=24}) + \frac{2(d\sum_{j=1}^{D}(m_{(j)}-1)+D)(k+1)}{d\sum_{j=1}^{D}(m_{(j)}-1)+D-k-2} \tag{2.9}$$

where d is the number of days required to ensure that $d\sum_{j=1}^{D}(m_{(j)}-1)+D > k+2$ and $\log(\sigma_{(j)}^2|_{t=i})$ is the value of $\log(\sigma_{(j)}^2)$ evaluated for the $t=i$ timepoint.

- <u>*Calculation of TI*</u>

The important aspect of Bayesian methodology necessary for model selection is the Bayes' Factor ratio of marginal posterior distributions [183, 184]:

$$B_{12} = \frac{p(j \mid M^{(1)})}{p(j \mid M^{(2)})} = \frac{\int p(j \mid M^{(1)}, \theta^{(1)}) p(\theta^{(1)} \mid M^{(1)}) d\theta^{(1)}}{\int p(j \mid M^{(2)}, \theta^{(2)}) p(\theta^{(2)} \mid M^{(2)}) d\theta^{(2)}} \quad (2.10)$$

where $p(j|M^{(1)})$ is the marginal likelihood that the data from dataset j was generated by model M^l, $p(j|M^{(1)},\theta^{(1)})$ is the likelihood of the data, j, being produced by model, $M^{(1)}$, with parameter set $\theta^{(1)}$ and $p(\theta^{(1)}|M^{(1)})$ is the prior probability distribution of the parameter set, $\theta^{(1)}$, being selected for model, $M^{(1)}$. B_{12} is the ratio of how likely one model is to represent the data compared against another model. The term I am therefore interested in is the probability $p(j|M^{(l)})$ – the larger this probability, the more likely the data was generated by that model, implying the model represents a simplified version of the true underlying biological mechanism.

Vyshemirsky & Girolami showed that one of the best approximations of $p(j|M^{(l)})$ is obtained using Thermodynamic Integration (TI) [183, 184]. For the purposes of this study, the problem shall be set out such that $j \sim \boldsymbol{N}(M^{(l)},\sigma^2)$ and $M^{(l)} \sim \boldsymbol{N}(\mu,\tau^2)$ is the prior distribution for our models, where $M^{(l)}$ is the time-series solution of the optimised model. The first approximation is that the data are normally distributed about the solution of the optimised model, with variance σ^2, where σ^2 is the same as in AICc above. The second approximation is that the solution of the optimised model is also normally distributed, with mean μ and variance $\tau^2 = \frac{\sum_{i=1}^{m} (M_i^{(l)} - \mu)^2}{m}$. In other words, τ^2 represents the divergence of simulations away from our prior knowledge of circadian waveforms, possibly due to parameter perturbation. To simplify the calculations, I am making the assumption that the prior model of circadian regulated transcription forms a sine-wave, providing rhythmic mRNA over a 24-hour day (see Figure 2.1). The relationship of perturbations in parameter space (chemotype) and function space (dynatype) has been described previously [203].

This is thus an empirically Bayesian technique, where the parameters of the Normal distribution for $M^{(l)}$, μ and τ^2, are approximated [204, 205]. As stated in the

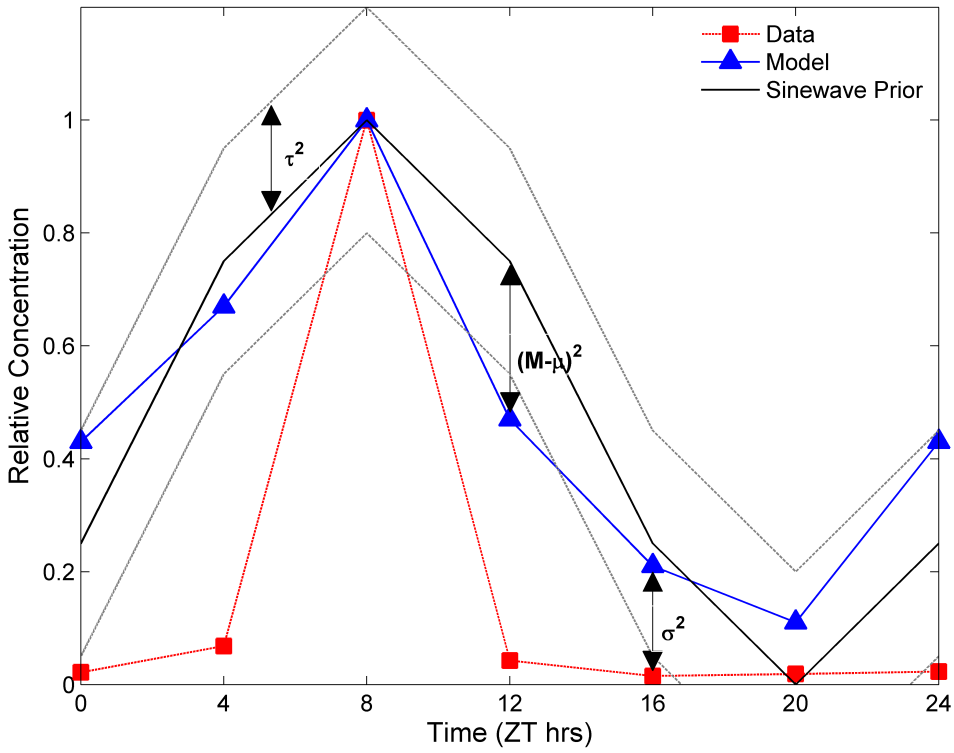


Figure 2.1: Generalised method used for comparing models. Models were fitted to data (red line, red filled squares) with the resulting example simulation (blue line, blue filled triangles). σ^2 is calculated as the sum of the differences between the model simulations and datapoints at the specific timepoints. A prior model was constructed using a sine-wave (black line). The difference between the models and prior model was calculated at the same timepoints as the models were compared to the data, $(M-\mu)^2$. The region around the prior model sine-wave (grey dashes) represents the space in which a model may lie given a perturbation to the prior model parameters. This space is characterised by the model variation, τ^2.

Introduction to this chapter, TI is generally used as part of parameter optimisation [183, 184]. However, the parameter values for the clock model are already known, so making the assumption that variation in the model is defined by a Normal distribution ensures that I do not have to directly calculate the variation in parameter values of the clock models.

In Bayesian statistics, to find out the marginal likelihood requires calculation of probabilities from a posterior distribution since

$$Posterior \propto Prior \times Likelihood$$

The posterior distribution in our case will be $M^{(l)}|j,\mu,\tau^2,\beta$ that tells us how likely the model $M^{(l)}$ is given all the information about uncertainty in parameter values. Following the example of Conditionally Independent Hierarchical Models from Kass & Steffey in deriving the posterior of Friel & Pettitt, the conditional posterior is

$$M^{(l)} \mid j,\mu,\tau^2,\beta \sim N\left(\frac{m_{(j)}\beta\tau^2\bar{j}+\sigma^2\mu}{m_{(j)}\beta\tau^2+\sigma^2},\frac{\sigma^2\tau^2}{m_{(j)}\beta\tau^2+\sigma^2}\right) \tag{2.11}$$

where β is the annealing temperature parameter of the distribution, $\bar{j}$ is the average value of the dataset and $m_{(j)}$ is the same as defined previously [183, 205]. Under normal Bayesian practices, sampling from this distribution would help determine the optimal model and parameter set in the region $\beta \in [0,1]$. This leads to the approximation for the marginal likelihood, which has the form

$$\mathrm{E}_{M^{(l)}|j,\beta}\left[\log\{p(j \mid M^{(l)})\}\right] = -\frac{m_{(j)}}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{m}(j_i-\bar{j})^2 - \frac{m_{(j)}}{2}\frac{\sigma^2(\mu-\bar{j})^2}{(m_{(j)}\beta\tau^2+\sigma^2)^2} - \frac{m_{(j)}}{2}\frac{\tau^2}{m_{(j)}\beta\tau^2+\sigma^2}, \tag{2.12}$$

where j_i is the same as $n_i^{(j)}$ previously. Using the trapezoidal rule as Friel & Pettitt with $\beta_z = (z/Q)^c$, where $Q = 40$ and $c = 4$ are arbitrary constants, approximations for $p(j|M^{(l)})$ are obtained as

$$\log\{p(j \mid M^{(l)})\} \approx \sum_{z=0}^{Q-1}(\beta_{z+1}-\beta_z)\frac{\mathrm{E}_{M^{(l)}|j,\beta_{z+1}}[\log\{p(j \mid M^{(l)})\}] + \mathrm{E}_{M^{(l)}|j,\beta_z}[\log\{p(j \mid M^{(l)})\}]}{2} \tag{2.13}$$

needed for model selection analysis, where z is the total number of discretized steps in $\beta \in [0,1]$ [183, 184]. The choice of Q and c are important in discretizing the temperature steps since the greatest change in $p(j|M^{(l)})$ takes place close to $\beta = 0$, hence having more steps near this point will lead to an approximation with greater accuracy than having $c = 1$ [183, 184].

When more than one dataset is used in the analysis

$$\mathrm{E}_{M^{(l)}|q,\beta}\left[\log\{p(q\,|\,M^{(l)})\}\right]=\sum_{j=1}^{D}\mathrm{E}_{M^{(l)}|j,\beta}\left[\log\{p(j\,|\,M^{(l)})\}\right] \tag{2.14}$$

where D and q are the same as the values defined previously. This assumes that the probability of one dataset being described by the model is independent of a second dataset being described by the same model. For future reference, $\mathrm{TI}_l = \log\{p(q|M^{(l)})\}$.

Model Selection Weights

For the AICc method of model selection, model selection weights (called Akaike Weights) have been discussed in previous studies (see [182]). Essentially they give a probability for a single model $M^{(l)}$ from the set of all l models. They take the form of

$$p(M^{(\ell)}) = \frac{\exp\{-\Delta^{\ell}/2\}}{\sum_{\ell}\exp\{-\Delta^{\ell}/2\}} \tag{2.15}$$

where $\Delta^l = \mathrm{AICc}_l - \mathrm{AICc}_{\min}$. This study uses the same technique for the TI analysis by taking $\Delta^l = -2(\mathrm{TI}_l - \mathrm{TI}_{\min})$. The reason for the factor of -2 in the case of TI 'weights' is that TI ~ log-likelihood, whereas AICc ~ -2*(log-likelihood). Thus $-\Delta^l/2$ ~ log-likelihood(l) – log-likelihood(min) in both cases. Due to this difference, the smallest AICc scores will be selected with the largest weight, whilst the highest TI scores will have the largest weights.

AICc with Uncertainty (AICc^U)

Since AICc does not include information about the prior probability of the models and TI does not directly penalise a model for over-complexity, I adapted the AICc to include model prior probabilities. To do this the AICc equation has been expanded to take into account prior information. Hence, from AICc

$$-2\log(L_{MLE}) = m\log(\sigma^2) + \log(2\pi\tau^2) + \frac{\sum_{i=1}^{m}(M_i^{(l)}-\mu)^2}{\tau^2} \tag{2.16}$$

where the term now includes the log-likelihood term for the prior distribution $M^{(l)} \sim N(\mu,\tau^2)$ from TI is now included (Figure 2.1). As can be seen from (2.16), results from AICc ≈ AICc^U when $\tau^2 \rightarrow \infty$ since the log(τ^2) term would become the same constant factor for each model. This would result in the same model being selected for AICc and AICc^U since $\Delta_{AICc} \approx \Delta_{AICcU}$. Similar ideas to this have been observed when considering the Bayesian Information Criterion (BIC) [182]. However BIC analysis assumes that the 'true' model is one of the models from the set being analysed, whereas AICc does not make this assumption and finds the 'best' model from the set – this results in different penalty terms being used between AICc^U and BIC.

TI with Parameter Penalty (TI^P)

In a similar manner to adding uncertainty to AICc, I wished to include the penalising of parameters in TI. To do this the penalty term from AICc was added to the values of TI, multiplied by a factor of -1/2 as with the TI 'weights' previously, such that

$$TI_l^P = TI_l - \frac{(d\sum_{j=1}^{D}(m_{(j)}-1)+D)(k+1)}{d\sum_{j=1}^{D}(m_{(j)}-1)+D-k-2}. \quad (2.17)$$

Computation

Optimisation of models was conducted using SBSI visual (http://www.sbsi.ed.ac.uk). Calculations and figures were obtained using Matlab (Mathworks Ltd, Cambridge, UK). All analysis was conducted using data from all of the photoperiods (SD, 12L:12D, LD) available.

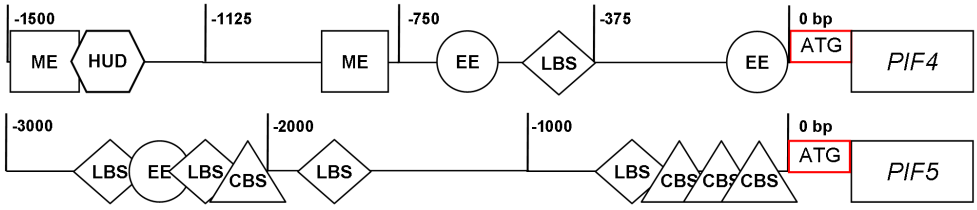


Figure 2.2: *cis* elements within the promoters of *PIF4* and *PIF5*. Schematic of the locations of consensus ME (AACCAC), EE (AATATCT), CBS (AAATCT), LBS (GATWCG) and HUD (CACATG) motifs in the promoter of *PIF4* and *PIF5* (defined as the 3kbp before the transcription start site, TSS, or to the nearest gene) found by manual sequence searches. For exact sequences see the 'Methods' section.

Results

Circadian regulated cis-elements are found in promoters of PIF4 and PIF5 genes

To determine potential circadian regulators of *PIF4* and *PIF5* transcription, I searched for previously identified, clock-regulated *cis*-regulatory sequences in the promoters of *PIF4* and *PIF5*. The Morning Element (ME), Evening Element (EE) and LUX binding site (LBS) are present in the *PIF4* promoter, whilst the CCA1 binding site (CBS), LBS and EE were found in the promoter region of *PIF5* (Figure 2.2). Interestingly, only the *PIF4* promoter has the Hormone Up At Dawn (HUD) binding domain, suggesting that it is this PIF that has transcription targeted by the hormones auxin and brassinosteroids and not *PIF5* [190]. Thus, across the promoters of both *PIF4* and *PIF5*, the presence of CCA1-targeted CBS and EE motifs (that lead to transcriptional activation and inhibition, respectively) and LBS motifs (that are targeted by the EC to inhibit transcription) allow us to consider models where the EC, CCA1 & LHY are candidate regulators of *PIF4* and *PIF5* transcription [19, 28]. The models were simplified by combining the PIFs together into one target gene, *PIF4/5*, in the same fashion as *CCA1* and *LHY* are combined as *CCA1/LHY* in the circadian clock models [12, 21, 23, 36–38].

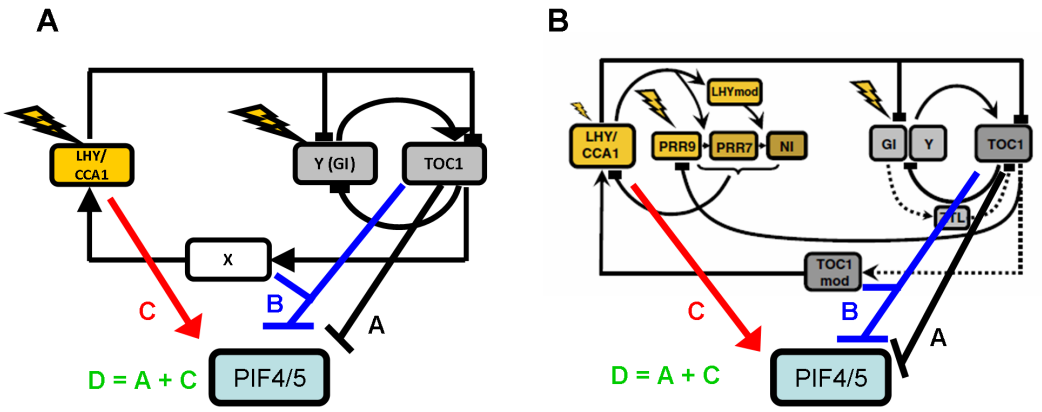


Figure 2.3: Model schematics of *PIF4/5* regulation. Four models of *PIF4/5* mRNA regulation by the circadian clock are highlighted using (A) the 'Locke2005' and (B) the 'Pokhilko2010' clock models. Model A (black) = Inhibition by TOC1 (referred to as Td); Model B (blue) = Inhibition by TOC1 and X (Td:Xd); Model C (red) = Activation by CCA1/LHY (Cu); Model D (green) = Model A + Model C = Inhibition by TOC1 and activation by CCA1/LHY (Td:Cu).

Rhythmic Data and Model Building

The promoter analysis above highlights that *PIF4* and *PIF5* transcription is regulated by the EC, as has been previously observed, and could potentially be regulated by CCA1 & LHY through the CBS and EE motifs (Figure 2.2; [25]). Combining this information with observations that *PIF4* and *PIF5* mRNA levels decrease in *cca1;lhy* mutants and increase in the absence of ELF3 and LUX function, I proposed that models of *PIF4/5* regulation required either inhibition by the EC and/or activation by CCA1/LHY [25, 178, 179]. However, neither circadian clock model used in this study explicitly includes a variable representing the EC (Figure 2.3). This meant that another clock component would need to serve as a proxy for EC activity. As the EC component *LUX* and *TOC1* mRNA are co-expressed and *toc1* mutants have a similar effect on *PIF4* and *PIF5* rhythm to *lux* mutants, then the existing TOC1 component of the clock models was used in place of LUX to achieve the requisite control of

PIF4/5 transcription in the evening (Figure 2.3; [23, 24, 38, 178]). Existing data in SD, 12L:12D and LD diel cycles (see Figure 2.4) showed that the rhythm of *PIF4* and *PIF5* transcript peaks in the morning with a trough at around dusk, while there is

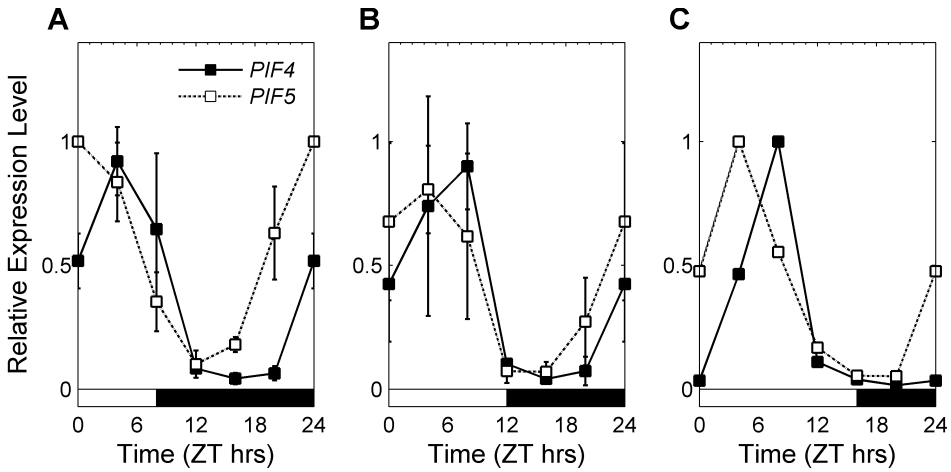


Figure 2.4: Transcription profiles of *PIF4* and *PIF5* used in this study. Data for *PIF4* (solid line, filled squares) and *PIF5* (dashed line, empty squares) mRNA in (A) SD, (B) 12L:12D and (C) LD cycles. Error bars represent standard deviation. No error bars shown in (C) due to lack of data.

a ~4hr phase difference between *PIF4* and *PIF5* mRNA (Figure 2.4). These observations are consistent with the proposition that *PIF4* and *PIF5* transcription is activated by CCA1 & LHY and inhibited by a representative evening repressor, such as the PRRs or EC [25, 31, 61, 178].

Interestingly, the analysis of promoter regions alone would not have highlighted the potential role of CCA1 & LHY promoting *PIF4* mRNA expression since the CBS is not present in the promoter. This suggests that either activation by CCA1 & LHY is indirect or that it occurs through the EE motif that is thought to be targeted to inhibit transcription. Alternatively, due to the absence of a CBS motif, inhibition of transcription by evening expressed components such as the EC (through the LBS) may be enough to sustain WT rhythms of *PIF4* mRNA with the required ~4hr phase delay compared to *PIF5* transcription. However, such a system would be unable to simulate the low amplitude transcriptional rhythms of *PIF4* observed in *elf3* and *lux* loss-of-function transgenic plants, suggesting that *PIF4* transcription is regulated by other circadian clock components [25, 179].

To examine whether *PIF4* and *PIF5* transcription is regulated by either the EC and/or CCA1 & LHY, mathematical model selection techniques can be used to

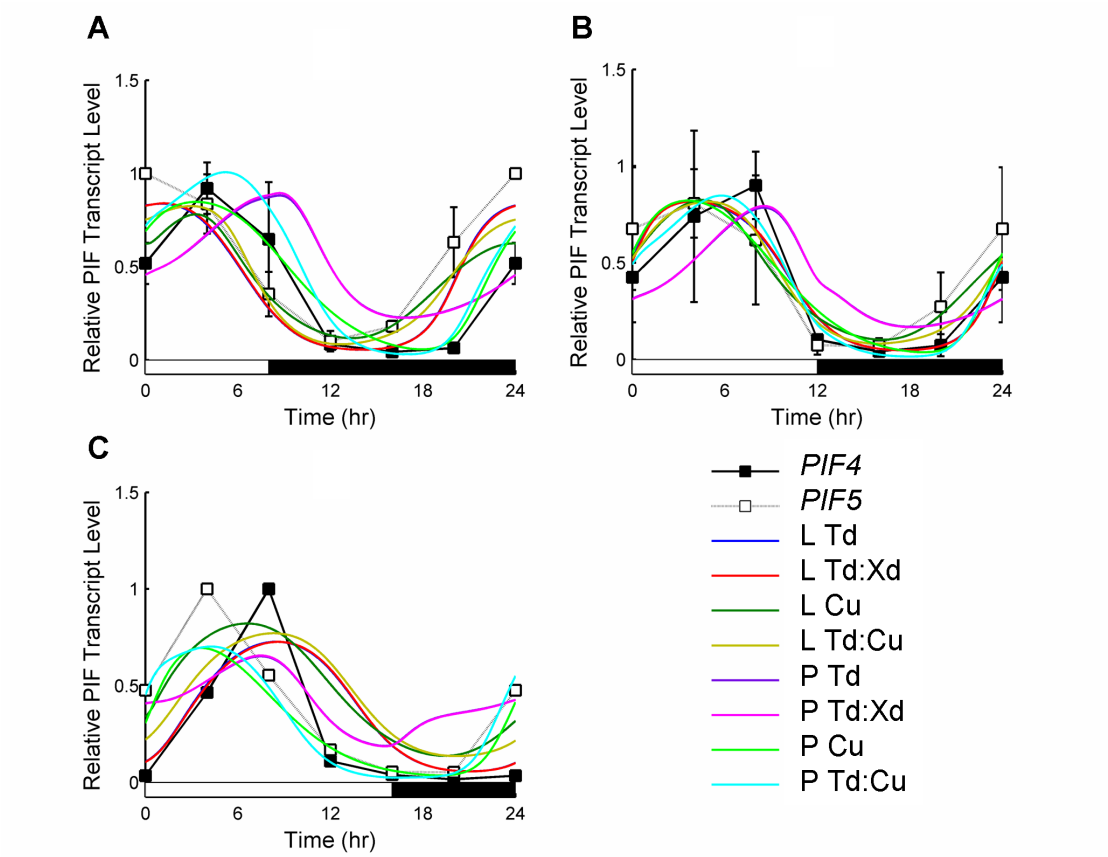


Figure 2.5: Simulations of *PIF4/5* mRNA using the models in Figure 2.3. Simulations of *PIF4/5* mRNA are shown in (A) SD, (B) 12L:12D and (C) LD and compared against data. *PIF4* data = solid black line, filled squares; *PIF5* data = dashed black line, empty squares. Simulations of *PIF4/5* in 'Locke2005' Td (L Td) model = blue line; L Td:Xd = red line; L Cu = green line; L Td:Cu = yellow line; 'Pokhilko2010' Td (P Td) = purple line; P Td:Xd = pink line; P Cu = bright green. Error bars represent standard deviation. No error bars shown in (C) due to lack of data.

determine how likely one system is compared to the set of analysed models. Thus, these techniques provide a method of comparing the performance of simple systems against complex models. In order to build a set of mathematical models describing *PIF4/5* mRNA, variants were constructed using either the 61-parameter 'Locke2005' or the more complex 90-parameter 'Pokhilko2010' circadian clock systems. Figure 2.3 shows schematics of the models tested in this study. The models feature either the 'Locke2005' or the 'Pokhilko2010' clock with *PIF4/5* being inhibited by TOC1 (referred to as Td in figures and tables); inhibited by TOC1 and X, an unknown within the clock model (Td:Xd); promoted by CCA1 (Cu), or; inhibited by TOC1

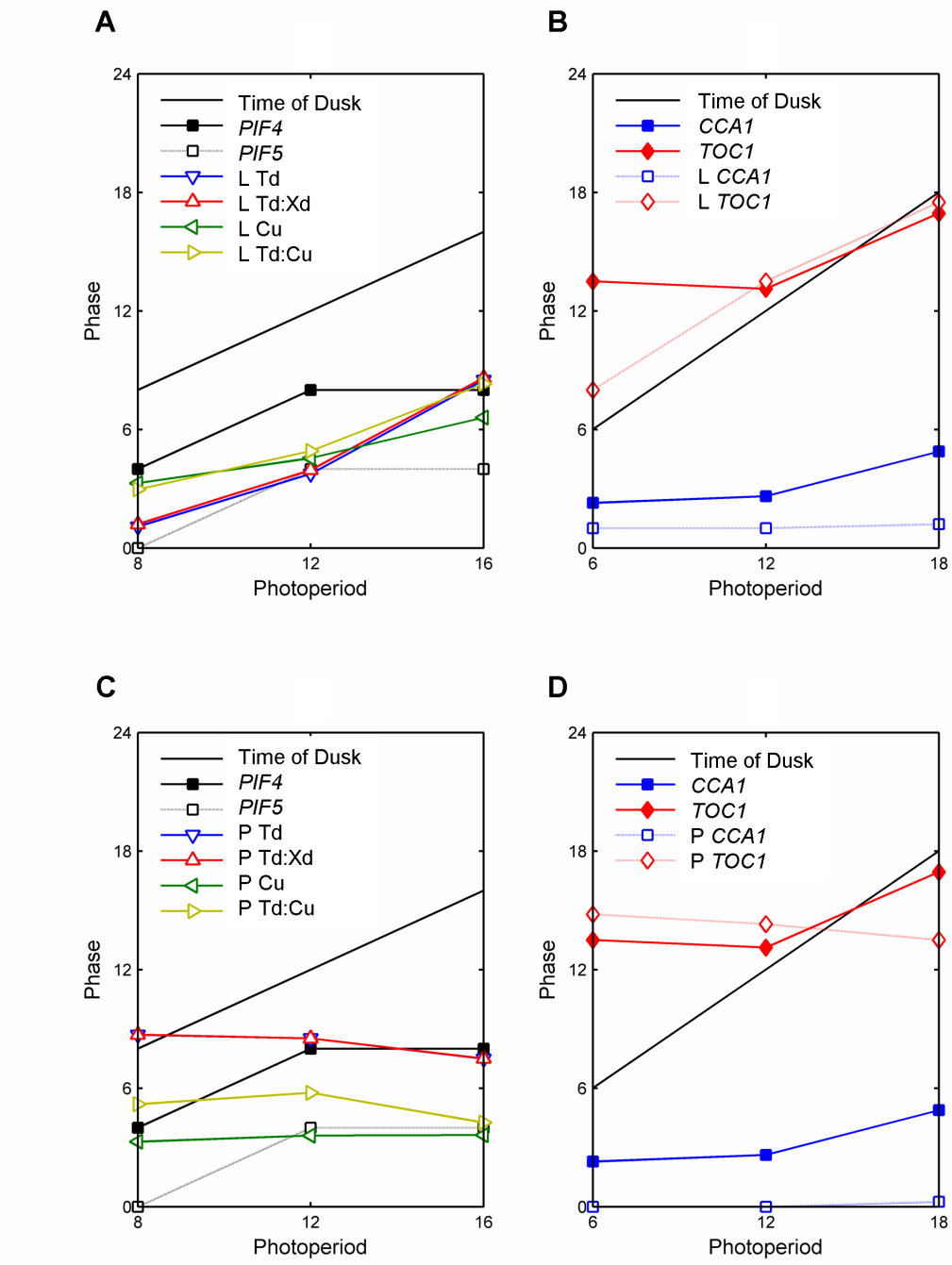


Figure 2.6: Comparison of relationship between photoperiod and phase of model components. Phase (time of peak expression) from simulations of (A) *PIF4/5* and (B) clock components *CCA1* and *TOC1* mRNA using models based on the 'Locke2005' circadian clock model are plotted against photoperiod of diurnal conditions simulated and compared with data. (C and D) Same as (A) and (B) for models based on the 'Pokhilko2010' circadian clock model. Time of dusk = black line; *PIF4* data = solid black line, filled squares; *PIF5* data = dashed black line,

and promoted by CCA1 (Td:Cu). The importance of testing the hypothetical component X as well as TOC1 is that X is expressed later into the night than TOC1 and was hypothesised to represent ELF4, a member of the EC [23, 25, 38]. This resulted in eight models of varying complexity and biological detail (Figure 2.3 andFigure 2.5; see Materials & Methods). As expected, simulations for *PIF4/5* mRNA have a similar relationship between phase (peak of expression) and photoperiod as the clock components that they are connected to (see Figure 2.6; [80]). Models featuring the 'Locke2005' clock had peaks of expression that shifted with the time of dusk as the photoperiod changed, whilst those that had the 'Pokhilko2010' clock tended to have more stable phases. As observed in Figure 2.6b and d, these results are expected as a consequence of the limitations of both clock models in describing data of clock genes in longer photoperiods. Interestingly, models that used the inhibition of *PIF4/5* transcription by TOC1 were almost identical to those that have *PIF4/5* mRNA inhibited by both TOC1 and X (Figure 2.5). This suggests that component X need not be considered in the system that regulates *PIF4/5* transcription.

Model Comparisons

- *Cost & Correlation*

Pearson correlation coefficients and a cost function (based on an averaged sum of square differences between the models and data; see Methods) represent two simple measures of how well a model fits data. Table 2.II shows the values obtained for each of the model simulations when compared to *PIF4* and *PIF5* mRNA time-series

---

(continued from above) empty squares; *CCA1* data = solid blue line, blue squares; *TOC1* data = solid red line, filled diamonds. Simulations of *CCA1* = dashed blue line, empty squares; *TOC1* = dashed red line, empty diamonds; *PIF4/5* using Td models = solid blue line, empty triangles; Td:Xd models = solid red line, empty triangles; Cu models = solid green line, empty triangles; Td:Cu models = solid yellow line, empty triangles.

Table 2.II: Cost and correlation scores for the models in Figure 2.3 excluding and including comparisons of *CCA1* and *TOC1* mRNA simulations with data from SD, 12L:12D and LD photoperiods. k = number of parameters.

		Excl. *CCA1* and *TOC1*		Incl. *CCA1* and *TOC1*	
Model	**k**	**Cost**	**Correlation**	**Cost**	**Correlation**
L - Td	65	0.26	0.73	1.28	0.7
L - Td:Xd	67	0.26	0.73	1.28	0.7
L - Cu	65	0.25	0.81	1.27	0.74
L - Td:Cu	67	0.24	0.76	1.26	0.72
P - Td	94	0.46	0.58	1.03	0.66
P - Td:Xd	96	0.47	0.58	1.03	0.66
P - Cu	94	0.26	0.78	0.83	0.76
P - Td:Cu	96	0.27	0.76	0.84	0.75

Table 2.III: AICc scores for the models in Figure 2.3 excluding and including comparisons of *CCA1* and *TOC1* mRNA simulations with data from SD, 12L:12D and LD photoperiods for varying values of *d*. k = number of parameters.

		Excl. *CCA1* and *TOC1*		Incl. *CCA1* and *TOC1*	
Model	**k**	***d*=2**	***d*=4**	***d*=2**	***d*=4**
L - Td	65	43.94%	0.03%	0.00%	0.00%
L - Td:Xd	67	0.43%	0.00%	0.00%	0.00%
L - Cu	65	5.25%	0.00%	0.00%	0.00%
L - Td:Cu	67	50.38%	99.96%	0.00%	0.00%
P - Td	94	0.00%	0.00%	0.00%	0.00%
P - Td:Xd	96	0.00%	0.00%	0.00%	0.00%
P - Cu	94	0.00%	0.00%	45.38%	1.30%
P - Td:Cu	96	0.00%	0.00%	54.62%	98.70%

Table 2.IV: TI scores for the models in Figure 2.3 excluding and including comparisons of *CCA1* and *TOC1* mRNA simulations with data from SD, 12L:12D and LD photoperiods. k = number of parameters.

Model	**k**	**Excl. *CCA1* and *TOC1***	**Incl. *CCA1* and *TOC1***
L - Td	65	28.12%	28.12%
L - Td:Xd	67	17.95%	17.95%
L - Cu	65	1.13%	1.13%
L - Td:Cu	67	52.79%	52.79%
P - Td	94	0.00%	0.00%
P - Td:Xd	96	0.00%	0.00%
P - Cu	94	0.00%	0.00%
P - Td:Cu	96	0.00%	0.00%

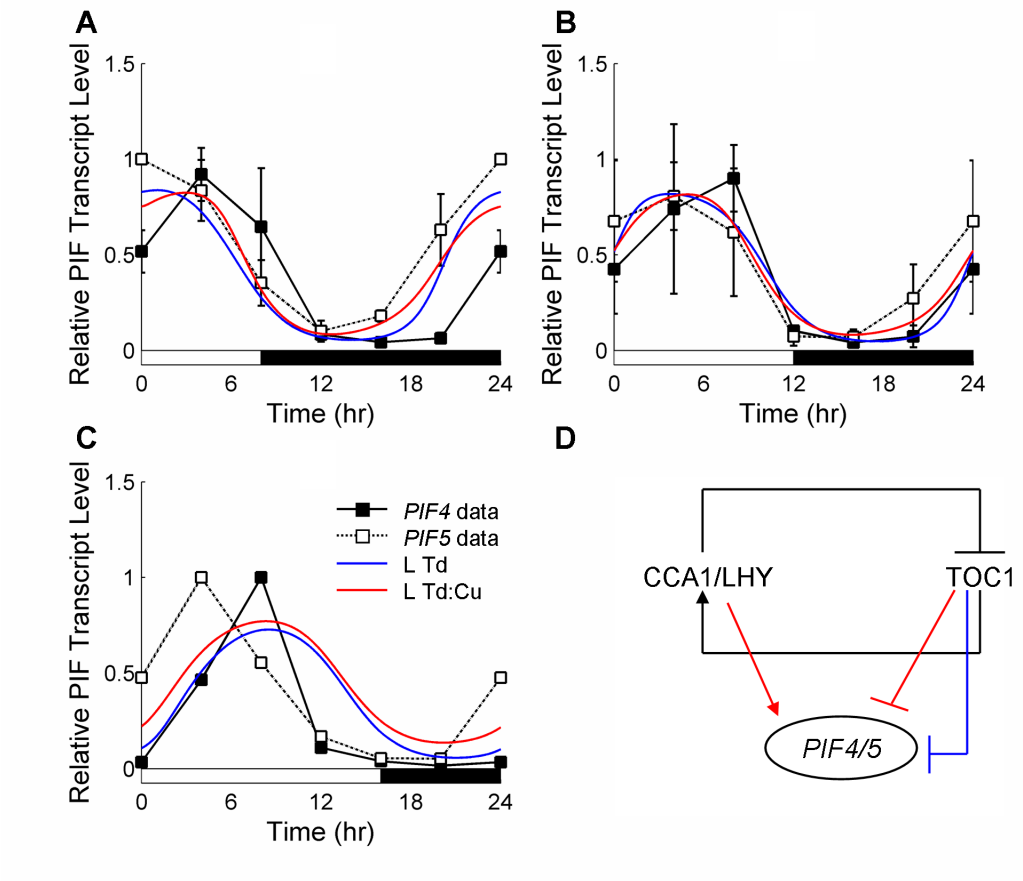


Figure 2.7: Simulations of *PIF4/5* mRNA from the models with the highest AICc and TI scores. Simulations of *PIF4/5* mRNA in (A) SD, (B) 12L:12D and (C) LD diurnal cycles using the 'Locke Td' (blue line) and 'Locke Td:Cu' (red line) models are compared to data. *PIF4* data = solid black line, filled squares; *PIF5* data = dashed black line, empty squares. (D) Schematic of the simulated models of *PIF4/5* transcription regulation by the circadian clock. Error bars represent standard deviation. No error bars shown in (C) due to lack of data.

from SD (8L:16D), 12L:12D and LD (16L:8D) photoperiods. From Table 2.II, a number of model simulations are observed to have similar cost and correlation values when models are compared to *PIF4* and *PIF5* expression data despite a difference in the number of parameters within the model. This implies that the added complexity from the 'Pokhilko2010' clock does not provide an advantage when trying to fit the model to data, suggesting that other model characteristics need to be considered in model selection analysis to differentiate between the two clock systems. However, if additional time-series data from clock transcripts (*CCA1* and *TOC1*) are included in the analysis then the 'Pokhilko2010' clock is clearly seen to be the more accurate model of both clock and output genes (Table 2.II). It is, however, important to

determine whether the improved cost score justifies the added complexity of 'Pokhilko2010'-based model variants.

- *AICc & TI*

Table 2.III shows that the most probable model for *PIF4/5* regulation found using AICc depends on the amount of data included in the analysis. When the data for *PIF4* and *PIF5* mRNA from the three photoperiods are considered alone then there is a 50.4% likelihood that the model variant selected to describe *PIF4/5* regulation is the 'Locke Td:Cu' model (for d = 2). Figure 2.7 shows how the most probable models of *PIF4/5* mRNA compare to the data sets of Figure 2.4. Including data of the clock components, *CCA1* and *TOC1*, across photoperiods leads to a 54.6% probability that these components and *PIF4/5* together are best described using the 'Pokhilko Td:Cu' model. This change of result is expected since the complex 'Pokhilko2010' model describes clock component mRNA profiles more accurately than the 'Locke2005' clock [23, 38, 80]. However, AICc does not account for any uncertainty in the function that simulates the model *PIF4/5* component.

The TI method takes into account the uncertainty of parameter estimation (as described in Methods). In contrast to AICc, the result of TI analysis was independent of the amount of genetic data included in the analysis (Table 2.IV). When *CCA1* and *TOC1* mRNA data is not considered, the model that most likely describes *PIF4/5* transcription in the photoperiods tested is the 'Locke Td:Cu' system (52.8% likely). Including *CCA1* and *TOC1* transcript data leads to the same result, with the 'Locke Td:Cu' model again being selected with a probability of 52.8% despite the models obtaining different TI scores (Table 2.IV). Hence, the regulation of *PIF4/5* mRNA from the 'Locke2005' clock by TOC1 inhibition and CCA1 activation is the model with highest probability regardless of the amount of information known. TI, though, does not penalise models for over-complexity, as AICc does. Thus, to test whether consistency could be obtained between the results of the AICc and TI analysis, I created hybrid techniques that take into account both a model's complexity and prior information about the models (see Methods).

- *$AICc^U$ & TI^P*

In order to consider both penalising a model for over-complexity and take into account prior information about the models, I combined the AICc and TI methodologies to form $AICc^U$ and TI^P (see Methods). Table 2.V shows the results of $AICc^U$ analysis and Table 2.VI shows the results of the TI^P model selection. Interestingly, adding approximate prior model distributions to AICc had no effect on the model selection, providing the same results as in Table 2.III: the 'Locke Td:Cu' model had a 53.5% weighting without clock data in the analysis, but the inclusion of clock data led to the 'Pokhilko Td:Cu' model having the highest likelihood (~61%). Penalising the TI analysis predictably provides more weight to the selection of the 'Locke Td' model, with its likelihood of selection increasing to ~94% (Table 2.IV). As with the original TI analysis above, where the 'Locke Td:Cu' model was selected regardless of how much data is included in the analysis, the 'Locke Td' model was again most likely when *CCA1* and *TOC1* transcript data was included in the analysis (85%). What this demonstrates is that these simple hybrid techniques are unable to provide consistent results between AICc and TI selection techniques. Possible reasons for this shall be outlined in the Discussion.

Effects of User-defined Variables

In the analysis outlined in the Methods section, there are two variables that could be defined by the user – the number of days of data used, d, and the size of the variance away from the model, τ^2. Here, results will be presented that show the effect of altering these values on the conclusions drawn from the analysis.

- *Changing the number of days, d, of data used alters the likelihood of a models selection*

As the total number of datapoints that was being used in the analysis was $\sum_{j=1}^{D}(m_{(j)}-1)+D=98$ and the largest model contained k = 96 parameters the

minimum value of d that could be used was d = 2 to ensure that $d\sum_{j=1}^{D}(m_{(j)}-1)+D>k+2$. As shown in Figure 2.8a, for all d used in the AICc and AICc^{U} analyses, the 'Locke Td:Cu' system was selected as the suitable model for *PIF4/5* transcriptional regulation from the set (the percentage values can be seen in Tables 2.III and 2.V for d = 4). However, as d was increased, the likelihood that the 'Locke Td:Cu' model was favoured from the set of models increased greatly. Studies of AICc analysis have previously recognised that if Δ^{l}>10 then the weight associatedwith the model would increase. Hence, to see whether this was why the probability increased, the difference between the penalty terms for the 'Locke Td:Cu' and 'Locke Td' models was calculated. As d increases, though, the difference in the penalty terms decreases to less than 10 (Figure 2.8b). This suggests that the large difference being generated between the AICc values for the two most likely models is due to an increased accuracy (or smaller cost) of the 'Locke Td:Cu' model. Therefore, the difference in the first term of (2.9) for these two models increases faster than the difference in the penalty term of (2.9) decreases. This would ensure that Δ^{l} > 10 for all d. The increased accuracy of the 'Locke Td:Cu' model explains why similar, albeit less significant, observations are made when the TI^{P} analysis is performed with d = *4*. In this situation, the likelihood of the 'Locke Td:Cu' model increases from 1.66% to 8.61%, leading to a decrease in the likelihood of the 'Locke Td' model from 93.99% to 85.04% (Table 2.VI).

- *Fixing all models to have the same variance, τ^2, can lead to poorly fitting models being favoured*

The other variable in the analysis that the user can define is the amount of variation allowed between the model and a standard sine-wave curve (used to represent our prior knowledge of a circadian rhythm, Figure 2.1). To determine how altering this value would effect the conclusions drawn from the analysis, AICc^{U}, TI and TI^{P} were

Table 2.V: $AICc^U$ scores for the models in Figure 2.3 excluding and including comparisons of *CCA1* and *TOC1* mRNA simulations with data from SD, 12L:12D and LD photoperiods for varying values of *d*. k = number of parameters.

		Excl. *CCA1* and *TOC1*		**Incl. *CCA1* and *TOC1***	
Model	**k**	***d*=2**	***d*=4**	***d*=2**	***d*=4**
L - Td	65	37.77%	0.03%	0.00%	0.00%
L - Td:Xd	67	0.37%	0.00%	0.00%	0.00%
L - Cu	65	8.35%	0.00%	0.00%	0.00%
L - Td:Cu	67	53.51%	99.97%	0.00%	0.00%
P - Td	94	0.00%	0.00%	0.00%	0.00%
P - Td:Xd	96	0.00%	0.00%	0.00%	0.00%
P - Cu	94	0.00%	0.00%	38.97%	1.00%
P - Td:Cu	96	0.00%	0.00%	61.03%	99.00%

Table 2.VI: TI^P scores for the models in Figure 2.3 excluding and including comparisons of *CCA1* and *TOC1* mRNA simulations with data from SD, 12L:12D and LD photoperiods for varying values of *d*. k = number of parameters.

		Excl. *CCA1* and *TOC1*		**Incl. *CCA1* and *TOC1***	
Model	**k**	***d*=2**	***d*=4**	***d*=2**	***d*=4**
L - Td	65	93.99%	85.04%	86.09%	79.34%
L - Td:Xd	67	0.56%	2.93%	2.65%	4.43%
L - Cu	65	3.78%	3.42%	3.46%	3.19%
L - Td:Cu	67	1.66%	8.61%	7.80%	13.04%
P - Td	94	0.00%	0.00%	0.00%	0.00%
P - Td:Xd	96	0.00%	0.00%	0.00%	0.00%
P - Cu	94	0.00%	0.00%	0.00%	0.00%
P - Td:Cu	96	0.00%	0.00%	0.00%	0.00%

performed with all models fixed to have the same value of τ^2 (see Figures 2.9 and 2.10). As observed in Figure 2.9a, over the range of τ^2 from 0.01 to 0.2, manually fixing a value of τ^2 for all of the models can provide different results to when τ^2 is calculated independently for each system analysed. For example, the 'Pokhilko Td:Cu' model would be the most likely model of *PIF4/5* mRNA if all the models were allowed the same small variation (τ^2<0.05). But, as the variation for all the models is steadily increased to 0.2, the weight of 'Pokhilko Td:Cu' against the other tested models decreases as the difference in $AICc^U$ values decreases. Similarly, the

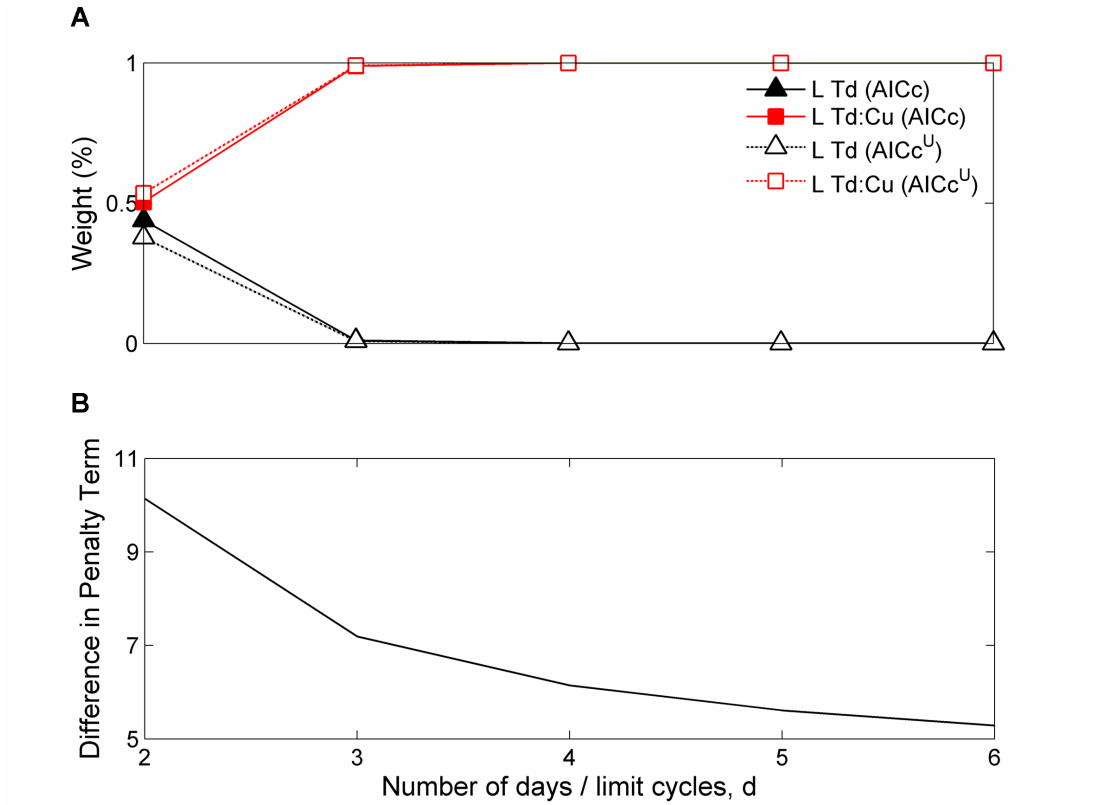


Figure 2.8: AICc and AICc^U scores are influenced by the number of limit cycles of data and simulations, *d*, are considered. (A) The Akaike Weights of the 'Locke Td' (black line) and 'Locke Td:Cu' (red line) from the AICc (solid line, filled squares) and the AICc^U (dashed line, empty squares) analysis are plotted against *d*. (B) The difference of the penalty term in the AICc analysis for the 'Locke Td' and 'Locke Td:Cu' models is plotted against *d*.

model with the highest probability from our initial analysis, the 'Locke Td:Cu' model, would only be recognized as the most likely model from the set if all the models were allowed a total variance of τ^2=16 from comparing simulations of *PIF4/5* in the three photoperiods to sine-wave curves (Figure 2.9b). However, having such a high variance in model dynamics would ultimately result in models with a poor fit compared to the data. The stars in Figure 2.9a highlight the values ofτ^2 that correspond to the AICc^U values obtained from the analysis, where all the models had independent values of τ^2, for the 'Locke Td' and 'Locke Td:Cu' models (see Methods). From the indicators, the smaller AICc^U score for the 'Locke Td:Cu' model corresponds with a slightly lower value of τ^2 compared to the 'Locke Td' model. Thus, models that are favoured by AICc^U tend to have smaller allowed variability from the sine-wave prior than other models in the set.

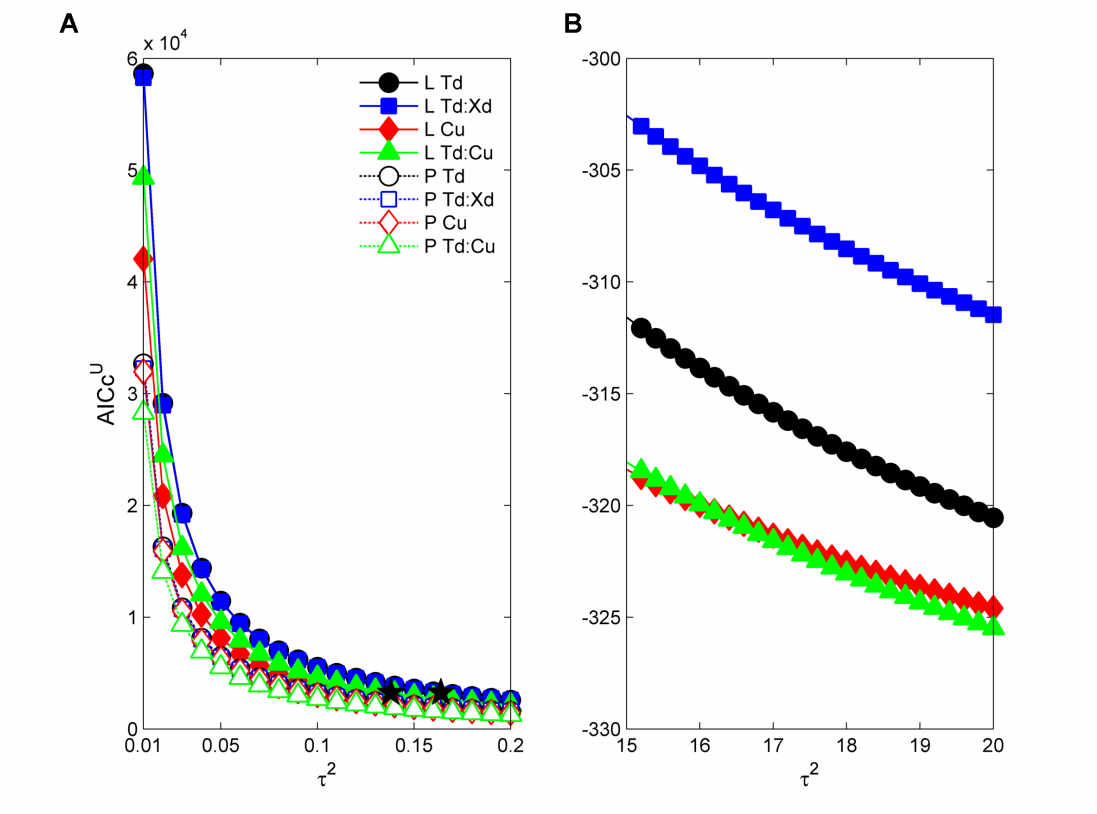


Figure 2.9: Values of AICc^U can change depending on the variance of the model parameters, τ^2. AICc^U values for the tested models were obtained for varying τ^2 with d = 2. (A) τ^2 = 0.01 to 0.2. (B) τ^2 = 15 to 20. 'Locke Td' model = solid black line, filled circles; 'Locke Td:Xd' = solid blue line, filled squares; 'Locke Cu' = solid red line, filled diamonds; 'Locke Td:Cu' = solid green line, filled triangles; 'Pokhilko Td' = dashed black line, empty circles; 'Pokhilko Td:Xd' = dashed blue line, empty squares; 'Pokhilko Cu' = dashed red line; empty diamonds; 'Pokhilko Td:Cu' = dashed green line, empty diamonds. Black filled stars in (A) indicate the AICc^U scores for the 'Locke Td' and 'Locke Td:Cu' models for default τ^2 values (see Methods).

Similar conclusions can be drawn by carrying out the same exercise with the TI and TI^P techniques (Figure 2.10). For τ^2<~0.18, the models found to be least likely by our initial TI analysis, the 'Pokhilko Td' and 'Pokhilko Td:Xd' models, would have the highest chance of selection even if the variance was fixed for all the models (as observed by high TI values). However, when τ^2>~0.18, these two models become the least likely, whereas the other models have a similar weighting (Figure 2.10a). When a penalty term is added in the TI^P analysis, the 'Pokhilko Td' and 'Pokhilko Td:Xd' models are again favoured if τ^2<~0.1 but the 'Locke2005' models are found to be

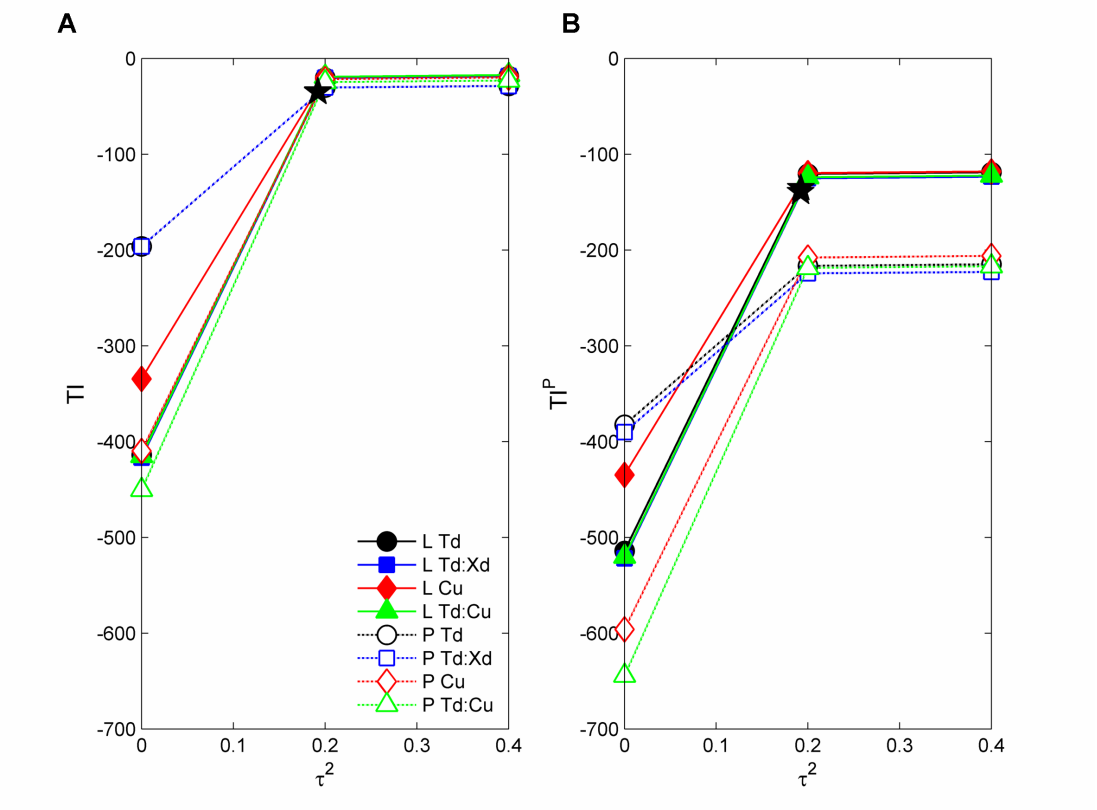


Figure 2.10: Values of TI and TI^P analysis with varying values of τ^2. As in Figure 2.9 for values obtained from the (A) TI and (B) TI^P analysis over a range of $\tau^2 = 0$ to 0.4. Black filled stars indicate TI and TI^P scores, respectively, for the 'Locke Cu' and 'Locke Td:Cu' models using default τ^2 values.

most probable when τ^2>0.1, which is the case in Table 2.VI (Figure 2.10b). Similarly to Figure 2.9, the stars in Figure 2.10a and b indicate the values of τ^2 that correspond to the obtained TI and TI^P scores from the analysis. In the TI analysis though, the stars highlight that having a slightly larger model variance can lead to a model with a larger TI score and higher likelihood of being selected (Figure 2.10). This analysis highlights two observations: first, allowing the estimated variance in each model to be calculated independently of each other removes the chance of poorly performing models to be favoured, and; second, models that are allowed a certain degree of variation away from a strict sine-wave function are more likely to be selected to describe a circadian rhythm across a range of conditions. However, having too much variation away from a sine-wave curve can lead to inconclusive results (as seen when τ^2>0.2 in Figure 2.10) or could lead to a model that has a poor fit with the data.

CCA1 has a stronger affinity for binding sites in the PIF5 promoter compared to exons in the PIF5 gene

Each of the analysis techniques above has supported models featuring activation of *PIF4/5* transcription by CCA1 and/or LHY, in addition to the known regulation of *PIF4/5* transcription by an evening repressor, such as the EC (Tables 2.III-VI; [25]). To test this model prediction a chromatin immunoprecipitation (ChIP) assay was conducted to test the binding of CCA1 to EE and CBS elements in the *PIF* promoters (Figure 2.2; D.R. MacGregor, University of Exeter, unpublished). As can be seen in Figure 2.11b, CCA1 was able to significantly bind to regions of the (control) *GI* promoter containing a CBS motif as has been published previously [179]. Furthermore, compared to an exon in the *PIF5* genomic region, CCA1 associated significantly with promoter regions of *PIF5* containing a CBS motif compared to exon regions in the *PIF5* gene, providing support that *PIF5* may be activated by CCA1 as proposed in this study (Figure 2.11). Interestingly, no significant enrichment of CCA1 binding was observed at *PIF4* promoter regions containing EE motifs. This suggests that either: CCA1 regulation of *PIF4* is much weaker than that of *PIF5*; *PIF4* is regulated by LHY, or; CCA1 activates *PIF4* expression indirectly, possibly through RVEs or PRRs [31, 46]. Each of these mechanisms should lead to the delayed phase of *PIF4* expression compared to *PIF5* whilst maintaining the high mean expression level, damp rhythms of *PIF4* transcription in *elf3* and *lux* mutants [25, 179].

Discussion

This study aimed to determine the most likely mechanism by which the circadian clock regulates expression of two growth promoting transcription factors, *PIF4* and *PIF5*. As well as matching the systems to experimental data, I wanted to examine the performance of models describing the circadian regulation of *PIF4* and *PIF5* transcription constructed using simpler variants of circadian clock models compared to newer, more complex systems. The calculation of simple cost functions and

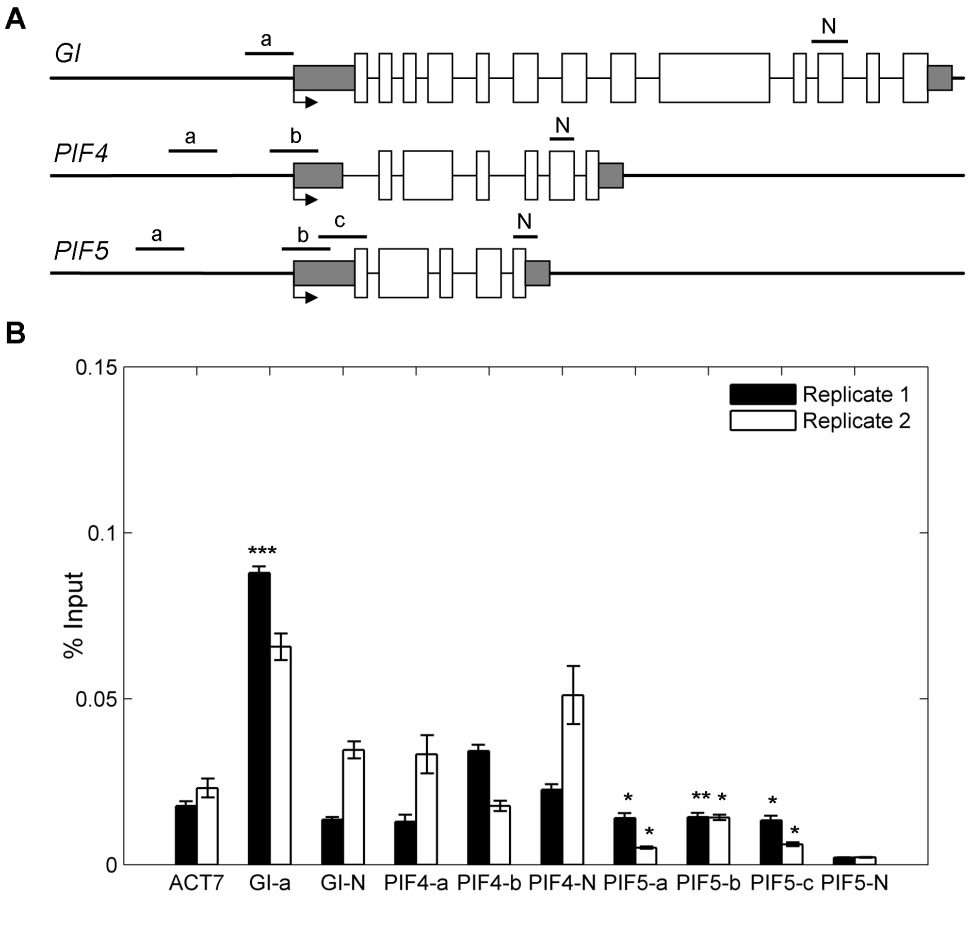


Figure 2.11: Validation of CCA1 binding to *PIF5* promoter. (A) Schematics of *PIF4* and *PIF5* promoters. *PIF4-a* and *PIF4-b* corresponds to regions of the promoter containing EE motifs (Figure 2.2). *PIF5-a*, *-b*, and *-c* correspond to regions of the promoter containing CBS motifs (Figure 2.2). Promoter regions *-N* are exon regions that are used as a negative control. Binding of CCA1 to the *GI* promoter is tested as a positive control [179]. *GI-a* contains a CBS motif. (B) ChIP analysis of CCA1 binding to the different regions for two biological replicates (D.R. MacGregor, University of Exeter, unpublished). Welch tests were used to test significant enrichment between promoter and exon regions: (*) $p < 0.1$; (**) $p < 0.05$; (***) $p < 0.005$. Error bars represent standard error.

Pearson correlation coefficients alone showed that the models of *PIF4/5* transcription could not be discriminated sufficiently to inform model selection, with 6 of the 8 models having similar cost and correlation scores (Table 2.II). Thus, despite the 'Pokhilko2010' clock model having 29 more parameters and 6 more components than the 'Locke2005' clock, this method did not indicate an improvement in the accuracy of *PIF4/5* mRNA simulations compared to the data. The lack of a significant difference in cost function meant that other model characteristics needed

to be taken into account to determine the most likely mechanism regulating *PIF4/5* mRNA.

Comparison of model selection techniques

Here, two statistical measures for model selection were used – AICc and TI. These methods take into account the complexity of the models (AICc) and the uncertainty associated with parameter optimisation (TI). These statistical measures were then expanded to test whether both model complexity and model uncertainty can be accounted for in hybrid techniques ($AICc^U$ and TI^P).

Whilst the AICc directly takes into account model complexity (as determined by the number of parameters in the system) to penalise the accuracy of simulations to data, this correction puts a constraint on the amount of data required to correctly use AICc analysis such that more data points are required than the number of parameters in the system (see Methods). In this chapter, the issue of limited data points was circumvented by repeating data from a single diurnal cycle over *d* days. Whilst this is used as an approximation of rhythmic transcription over a number of days, this simplification neglects changes in transcriptional rhythms that may occur from day to day. For example, Nozue and co-workers have shown that, under diurnal cycles, mRNA rhythms of *PIF4* and *PIF5* appear to have altered mean expression levels from one day to the next. Furthermore, daily rhythms of hypocotyl elongation dampen over a 3-day period [61]. Thus, a potentially better way of increasing the number of data points for an analysis such as this would be to increase the resolution of mRNA levels over a single diurnal cycle. This would allow for the AICc analysis to be performed without the need of including parameter *d*.

Normally, the TI method is used as part of parameter optimisation, helping to find the most likely model and parameter set that describes experimental data. Consequently, for large and complex models, methods such as TI may require large computational power and take a long period of time to effectively explore all the dimensions of parameter space to find optimal parameters for a system. Here, TI was

used to calculate a log-likelihood probability that would help with model selection analysis, similar to AICc. In order to account for the divergence of potential models from a 'prior' model in the selection techniques I assumed that any variation in parameter value would result in a normal distribution of possible models around a standard sine-wave curve that acts as a first approximation (or prior knowledge) of circadian regulated transcription. This assumption means that this methodology can be referred to as an empirical Bayesian technique (see [206]). True Bayesian techniques do not make assumptions of this nature, calculating complex probability distributions for parameter values as part of the model optimisation process. The empirical Bayesian techniques therefore offer an approximation for these distributions whilst also cutting down on computational cost.

The initial results of the analysis favoured the 'Locke Td' and 'Locke Td:Cu' models (Tables 2.III and 2.IV). This supports the hypotheses drawn from analysis of the *PIF* promoters and mRNA time-series from *elf3*, *lux* and *prr9;7;5* loss-of-function transgenic plants that highlighted a potential role for CCA1 & LHY in the regulation of *PIF4* and *PIF5* transcription conducted as part of this study [25, 178, 179]. Furthermore, this confirms the previously published observations showing that an evening repressor that is co-expressed with TOC1 (known to be the EC or other PRRs) is necessary to regulate *PIF4* and *PIF5* transcription (Figure 2.2; [25, 31]). To explore the role of CCA1 & LHY in *PIF4* and *PIF5* mRNA regulation, ChIP analysis showed that CCA1 associated with promoter regions of *PIF5* that contained CBS motifs compared to exon regions in the *PIF5* gene (Figure 2.11; D.R. MacGregor, University of Exeter, unpublished). However, as the signal of CCA1 binding to the *PIF5* gene is lower than that of CCA1 binding to the *ACT7* gene used as a negative control it is difficult to confirm with confidence that CCA1 directly regulates *PIF5* transcription. Similarly, the ChIP assay showed no evidence of CCA1 enrichment at the *PIF4* EE promoter motifs compared to the *PIF4* exon or the *ACT7* negative control suggesting that CCA1 may be more likely to activate *PIF4* (and *PIF5*) transcription indirectly, potentially through the RVEs that target EE motifs and activate transcription or the PRR protein family that inhibit transcription [31, 46]. Whilst the models presented in this study (with a multiplicative positive term of

PIF4/5 mRNA regulation by CCA1 protein) would not be able to accurately describe the rhythmic expression of *PIF4* and *PIF5* observed in *cca1;lhy* loss-of-function mutants. Rhythms of *PIF4* and *PIF5* transcript levels have lower mean expression levels and maintain a strong peak at around ~ZT0-3 in *cca1;lhy* transgenic plants [178]. But, if $c_{CCA1/LHY}$ (or c_A) in (2.1) is set to zero, which would be the case for a simulation describing the *cca1;lhy* loss-of-function mutant, then $c_{PIF4/5}$ would not change with time ($dc_{PIF4/5}/dt = 0$) implying that *PIF4/5* transcription would not be rhythmic. This highlights a minor role for a regulator of *PIF4/5* transcription that is independent of CCA1 & LHY (see Chapter 4 and Appendix B).

Interestingly, the results shown in Tables 2.III – 2.VI may highlight cases when it is appropriate to use AICc model selection techniques over TI-based techniques. The AICc-based analyses found that the 'Locke Td:Cu' model was the most likely system to correctly describe *PIF4/5* mRNA across photoperiods from the set of models tested (Tables 2.III and 2.V). Additionally, when clock data is included in the analysis, the AICc model selection techniques favour models constructed using the 'Pokhilko2010' clock that describes circadian dynamics with higher accuracy than the 'Locke2005' clock model (Tables 2.III and 2.V). Therefore, AICc-based model selection techniques appear to be more useful to aid construction of mathematical models in cases where the circadian clock dynamics are important for the system. This would be an important factor in model construction when one wishes to describe genetic or environmental perturbations to the clock machinery that cannot be described by simple clock systems. Conversely, as seen in Tables 2.IV and 2.VI, TI-based model selection techniques provide the same result independent of the inclusion of information about clock dynamics (in the examples presented here, this is by including *CCA1* and *TOC1* mRNA rhythms). This would be useful for cases requiring simple systems that are able to describe WT diurnal rhythms across a range of photoperiods. These simple systems can then be used as modules in larger models to describe plant physiology [82]. Therefore, an important part of model construction is determining the level of detail that is required for the model to capture. Once this decision has been made the model selection process can begin.

Increased amount of data and parameter variation alters conclusions of model selection techniques

From the various model selection techniques used in this study, two key observations were made. First, the conclusions that were drawn from the analysis changed depending on the data that was considered. During this study I increased the amount of data used in the analysis in two ways. Initially, by including data for *CCA1* and *TOC1* mRNA that both the 'Locke2005' and 'Pokhilko2010' models should be able to describe resulted in all of AICc-based techniques favouring models based on the 'Pokhilko2010' clock (Tables 2.III-VI). This is due to the improved accuracy that the newer and more complex clock models have over simpler model variants when comparing data of clock components. However, TI-based techniques chose the same models regardless of the amount of information included in the analysis ('Locke Td:Cu' in TI and 'Locke Td' in TI^P) as the 'Locke2005' circadian clock is favoured over the 'Pokhilko2010' system using TI. This suggests that, as well as sufficiently matching the data, the 'Locke2005' clock model also has simulations with a closer match to the 'prior' sine-wave functions compared to the 'Pokhilko2010' system. Thus, even the improved cost of the 'Pokhilko2010' clock compared to the 'Locke2005' clock, as seen in Table 2.II, is unable to improve the likelihood of 'Pokhilko2010' based models using TI.

The second way in which I was able to alter the amount of data used in the analysis was by changing the value d. This led to an increase in the number of time-points used in the analysis by repeating a single 24-hour rhythm over d days. In doing this, I found that the 'Locke Td:Cu' gained increasing support over the other *PIF4/5* models (Figure 2.8). This result is due to the 'Locke Td:Cu' model being more accurate than other model variants leading to a lower AICc score and this difference becomes extrapolated as the number of limit cycles/days is increased. As the AICc has an approximately linear relationship with d (i.e., AICc ~ $Ad + B$, where A is the models accuracy to data such that higher values imply less accuracy and B is the penalty term; compare to Eq. (2.9)), a higher value of A would lead to a faster increase of AICc score as d is increased compared to a lower value of A.

Mathematically, this can be written as: if $A_1 > A_2$, leading to two AICc scores $C_1 = A_1d + B_1$ and $C_2 = A_2d + B_2$, then $dC_1/d(d) = A_1 > A_2 = dC_2/d(d)$, thus explaining why the AICc score C_1 would increase faster than C_2 with increased values of d resulting in a higher AICc weight for model C_2. A similar, albeit weaker, observation was made as a result of the TI analysis with increasing d, whereby the 'Locke Td:Cu' gained increased weighting due to improved model accuracy over the 'Locke Td' system (Table 2.VI).

The increases in data led to the AICc-based model selection techniques favouring more complex systems, whereas the TI-based methods favoured the same model regardless of the amount of data used. The reason for this difference between AICc- and TI-based techniques is the different weighting given to simulation comparisons with data and the prior sine-wave model in the calculation of AICc and TI scores. The AICc-based methods weight the accuracy of models to data to a higher degree than TI-based methods, which also consider comparisons between 'prior' and resulting models. Thus, whilst simulations of *CCA1/LHY* and *TOC1* mRNA by the 'Pokhilko2010' model are more accurate than simulations from the 'Locke2005' system (leading to higher AICc weights for 'Pokhilko2010'-based model variants when this data is considered in the analysis), the increased complexity results in simulations that diverge to a greater degree from the simple sine-wave 'prior' model of circadian transcription (leading to 'Pokhilko2010'-based model variants having lower TI weights than 'Locke2005' variants).

The second observation made was that the model selection techniques are sensitive to the amount of variance a model is permitted to have away from a standard sine-wave curve. The value of τ^2 represents the variance of the model simulations away from the 'prior' sine-wave model. By manually fixing the value of τ^2, I was able to calculate $AICc^U$, TI and TI^P scores for each *PIF4/5* model. As seen in Figure 2.9 and 2.10, if all the models were forced to have small values of τ^2 (less than ~0.18 in the case of TI and TI^P and less than ~0.05 for $AICc^U$), the analysis would produce inconsistent results. For example, using $AICc^U$, the more complex 'Pokhilko Td:Cu' model would be found to be more likely than a number of other, simpler systems

(Figure 2.9a). Similarly, using TI and TI^P, the two models with the highest cost score, 'Pokhilko Td' and 'Pokhilko Td:Xd', would be favoured ahead of much better data-fitting models (Figure 2.10 and Table 2.II). This analysis suggests that whilst having too much variance in the models away from a standard sine-wave curve would lead to poorly fitting models, having too little variance can lead to models that are more complex or poorly fitting to be favoured by the analysis techniques presented here.

Conclusion

What I have shown in this chapter is that when a models accuracy to data (or cost) is penalized for parameter numbers or by the models similarity to a standard 'prior' function, simpler models are favoured over those with more complexity. However, more elaborate models, if constrained sufficiently by data, provide a better realisation of the biological system. For example, it has already been noted in this chapter, that the models presented here would be unable to correctly characterize *PIF4/5* transcription in the case of a genetic double mutation of *cca1;lhy*. Including a condition in the techniques presented above to select a model that can not only describe *PIF4/5* transcription in different photoperiods but also in different genetic perturbations may yield results favouring more complex networks. However, in such a case, defining a suitable prior model to describe rhythms in transgenic plants may prove difficult.

Furthermore, the sensitivity of the techniques to model variance shows that care needs to be taken when placing specifications on how much variation is allowed in parameter values. If there is too much variation, then the model becomes inaccurate. However, if there is very little variation then complex models out-perform the simpler systems. What is clear is that the number of parameters used in a system and the range of values that those parameters can take should be constrained. The most sensible way of limiting a model would be by the amount of variation observed in the experimental data. Hence, the variation in a model would not be greater than what is

seen in real biological systems. This limiting effect should provide a balance between describing the data simply whilst maintaining accuracy.

Through the rest of this thesis, these views will be taken into account when developing mathematical models for the system that controls photoperiodic flowering in *Arabidopsis thaliana* and how light and temperature interacts with these pathways.

Chapter 3: Mathematical Modelling of the Photoperiod-dependent Flowering System in *Arabidopsis thaliana*

This work was published as part of:
Y. H. Song, *et al.* (2012) FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science*, vol. 336, pp. 1045-9.

In the previous chapter, part of the philosophy of mathematical modelling was discussed whereby simple models can often provide more useful information about a systems dynamics than more complex models. However, this depends on the question that is being asked of the model. In this chapter, a model of the flowering pathway in *Arabidopsis thaliana* will be built with the explicit purpose to quantify how much of a role the blue light-regulated protein, FKF1, plays in the system. As such, a more complex variant of the photoperiodic flowering model is required to examine the functions of FKF1 in detail.

Previous model of Arabidopsis flowering predicts an FKF1-dependent feed-forward network

As discussed in Chapter 1, a model describing the external coincidence flowering system of Arabidopsis has previously been constructed [82]. In this model, the circadian clock and FKF1, a blue light-regulated protein, activates *CO* transcription prior to dusk in LDs [74, 82]. This leads to an FKF1-dependent 'shoulder' in *CO* expression at ~ZT13 of LDs (Figure 3.1a; [74]). The 'shoulder' is not observed in SDs as FKF1 levels coincide with darkness, when the protein is unable to perform its blue light-dependent functions [82]. This is known as external coincidence (see Chapter 1; [79]). The CO protein, in turn, promotes transcription of *FT* such that the expression levels of *FT* are increased in LDs relative to SDs [73, 75]. However, the simulations of *fkf1* mutants showed strong rhythms of *FT* mRNA in LDs that are not present in *fkf1* loss-of-function transgenic plants, where *FT* levels are basal

throughout the diurnal cycle [82]. Salazar and co-workers, therefore, derived a model-based hypothesis that FKF1 must act downstream of *CO* mRNA to regulate *FT* transcription in a CO-dependent manner [82].

In addition to *FT* transcription being controlled by multiple components, recent evidence has shown that the regulation of *CO* transcription by FKF1 is more complex than described by this simplified model. Transcription of *CO* has been shown to be indirectly regulated by FKF1. FKF1 protein forms a blue light-dependent protein complex with the circadian clock component GI [81]. Through the F-box domain, FKF1 can also form an SCF ligase complex that targets proteins for degradation [102]. Consequently, the GI-FKF1 protein complex has been found to degrade inhibitors of *CO* transcription [83, 90]. The CDF protein family suppresses *CO* mRNA levels throughout the day such that constitutive overexpression of CDF1 protein leads to low levels of *CO* expression [90]. Levels of CDF1 & 2 protein are dependent on GI and FKF1 [83, 90]. For example, the loss of FKF1 function in transgenic plants results in the slower degradation of CDF1 during the second half of long photoperiods [90]. Thus, the apparent activation of *CO* mRNA by FKF1 occurs due to GI- and FKF1-dependent inhibition of CDF protein that, in turn, suppresses *CO* expression.

In this chapter, the model of flowering will be expanded to accommodate the post-translational regulation of CDF1 (as a representative of the CDF protein family) alongside evidence provided by collaborators that validates an interaction between CO and FKF1 proteins. The resulting model aims to answer the following question: relatively, how important is the interaction of CO-FKF1 compared to the interaction of CDF1-FKF1 to regulate transcription of *FT*?

Methods

Data analysis for modelling

Data acquired for this study was obtained by Dr. Young Hun Song, Dr. Benjamin To and Assistant Prof. Takato Imaizumi (University of Washington, USA). This has since been published in [153].

To maintain consistency with the previous model, the same WT and *fkf1* data was used to optimise the model parameters describing *CO* and *FT* mRNA [82]. These datasets were labelled as co1, co3, co8, co9, ft1, ft3, ft8, ft9, co8fkf1, co9fkf1, ft8fkf1 and ft9fkf1 in [82]. Datasets co1, co3, ft1 and ft3 represent SD and LD *CO* and *FT* mRNA from [76]. Datasets co8, co9, ft8 and ft9 are LD and SD *CO* and *FT* mRNA in WT background and co8fkf1, co9fkf1, ft8fkf1 and ft9fkf1 are the same in the *fkf1* background from [74]. The newly-obtained data for *CO* and *FT* mRNA in wild-type plants were used to validate the new model. The new WT control data are qualitatively similar to the older data sets showing that experimental variation does not greatly affect the experimental results for *CO* and *FT* mRNA (Figure 3.1; [153]). This provides justification for testing the model against data acquired across different experiments from mutant backgrounds. Newly obtained data was normalized in a similar manner to [82] such that the SD peak of *CO* mRNA and the LD peak of *FT* mRNA in WT conditions were set to 1. All data was then normalized against a common internal standard, the maximum level of SD *CO* mRNA from dataset co9 that was set to 1.

Further to the *CO* and *FT* mRNA waveforms obtained in this study, *CDF1* mRNA and protein levels were measured. *CDF1* mRNA and protein has been shown to have a circadian rhythm (Figure 3.2; [50, 78, 83, 90]). The circadian regulation of *CDF1* mRNA will be discussed in more detail in Chapter 4. Therefore, rather than introducing hypothetical regulators and unknown parameters, experimental data of *CDF1* mRNA is used as an input to the model as was the case with FKF1 protein in the previous flowering model [82]. Data for *CDF1* mRNA was acquired in WT, *fkf1*,

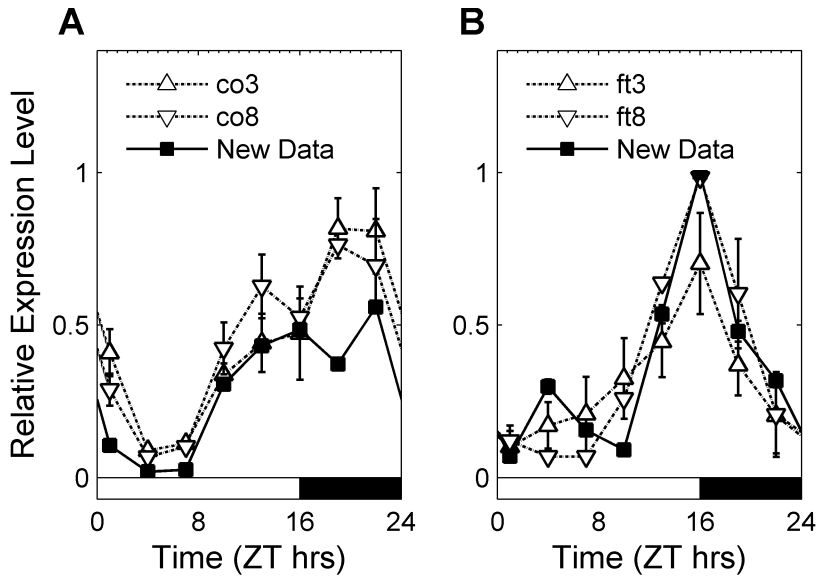


Figure 3.1: Comparison of *CO* and *FT* mRNA from a range of experiments. Expression of (A) *CO* and (B) *FT* mRNA were obtained from three LD experiments; co/ft3 = dashed black line, empty upright triangles [76], co/ft8 = dashed black line, empty downturned triangles [74] and new data generated for this study = solid black line, filled squares [153]. Error bars represent standard error.

35S:CO, *35S:CO;fkf1* and*35S:CDF1;SUC2:CO* lines (Figure 3.2a-c), whereas CDF1 protein levels were obtained in WT and *fkf1* plants (Figure 3.2d). To maintain simplicity, FKF1 protein was maintained as a direct input into the model to avoid adding hypothetical regulators of *FKF1* mRNA and protein.

The waveform of *CDF1* mRNA was normalized to the LD peak in the WT background. Other than in the *35S:CDF1* background, the *CDF1* waveform was relatively unaltered in all other backgrounds, allowing us to use WT *CDF1* mRNA as the system input for the majority of tested conditions (Figure 3.2a-c). As the protein data for CDF1 and FKF1 were comparable, all protein data was normalized to the LD peak of FKF1 in the WT background (Figure 3.3). Similarly, as data was obtained that allowed for the comparison of CDF1 protein in WT and *fkf1* backgrounds, the difference in protein level between these two backgrounds was estimated (Figure 3.2d). The normalized data were used to constrain the model, as described below.

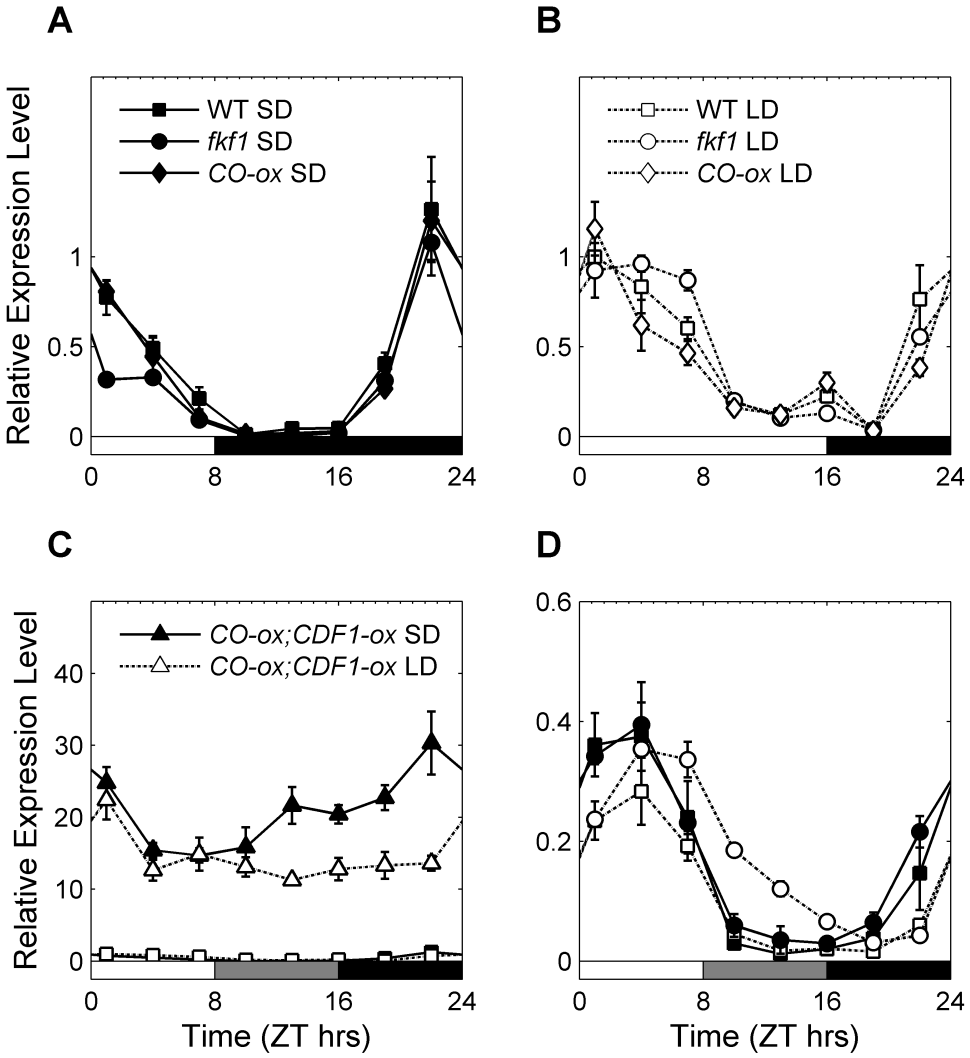


Figure 3.2: Relative expression levels of *CDF1* mRNA and CDF1 protein. *CDF1* mRNA expression was obtained for this study in (A,B) WT (squares), *fkf1* mutant (circles), *SUC2:CO* (*CO-ox*; diamonds) and (C) *35S:CDF1;SUC2:CO* (*CO-ox;CDF1-ox*; triangles) overexpressors in both (A, C) SD (solid line, filled shapes) and (B, C) LD conditions (dashed line, empty shapes). (D) CDF1 protein levels were obtained from WT and *fkf1* mutants in both SD and LD conditions. Error bars represent standard error.

Model derivation

- *Revised connection of the circadian clock to CO transcription*

As described in the introduction to this chapter, the previous model of flowering time used direct activation of *CO* transcription by the GI-FKF1 protein complex rather than the repression of *CO* mRNA by CDF1 that is, in turn, degraded in a GI-FKF1

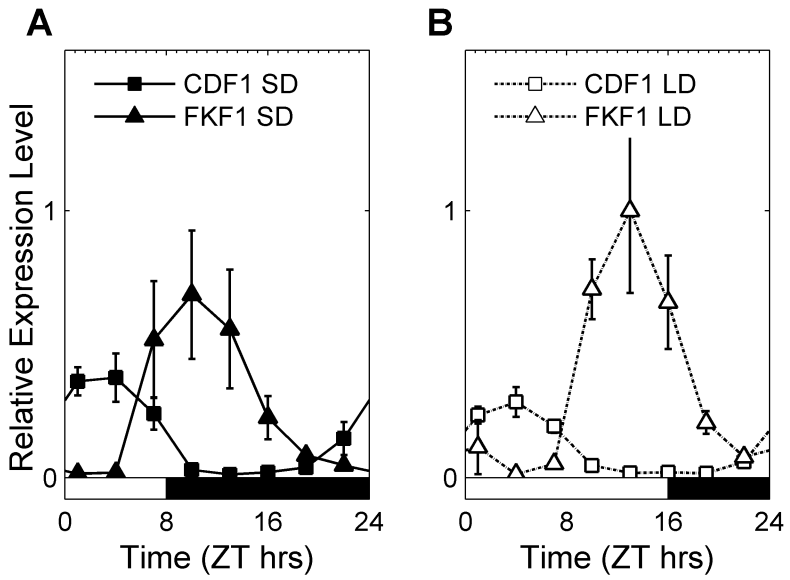


Figure 3.3: Comparison of protein levels of FKF1 and CDF1. By obtaining relative amounts of FKF1 (triangles) and CDF1 (squares) protein, the difference in expression was estimated in both (A) SD (solid line, filled shapes) and (B) LD conditions (dashed line, empty shapes). Error bars represent standard error.

dependent manner [81, 82, 90]. Using *CDF1* mRNA and FKF1 protein as an input to the system, the model is able to simulate CDF1 protein levels (see Figure 3.8a and b in Results section below). This system, though, would still be unable to describe the mechanism that generates the LD specific night-time peak of *CO* mRNA that occurs independently of FKF1 (see Figure 3.8c and d in the Results section below). Hence, in a similar manner to the previous model, I used the 'Locke2005' clock model where *TOC1* mRNA tracks dusk and is a good fit for the hypothetical component that is an additional regulator of *CO* transcription in LDs [23, 82]. The 'Locke2005' clock also has the added characteristic that the system comprises GI explicitly, allowing us to model the light-dependent GI-FKF1 effects on CDF1 protein. These procedures allow the model to simulate *gi* and *cdf* mutants (as a proxy for the *cdf1-R;2;3;5* mutant; Figure 3.4; [81, 83]). The qualitative differences of the *gi;cdf* (representing *gi;cdf1-R;2;3;5*) simulations with the data suggests that the presented model does not accurately capture the relationship of GI with CDF proteins (Figure 3.4). However, since this model has been constructed to assess the role of FKF1 in the flowering system, the differences between *gi;cdf* simulations and data shall not be examined here (see Chapter 4). The model is shown schematically in Figure 3.5.

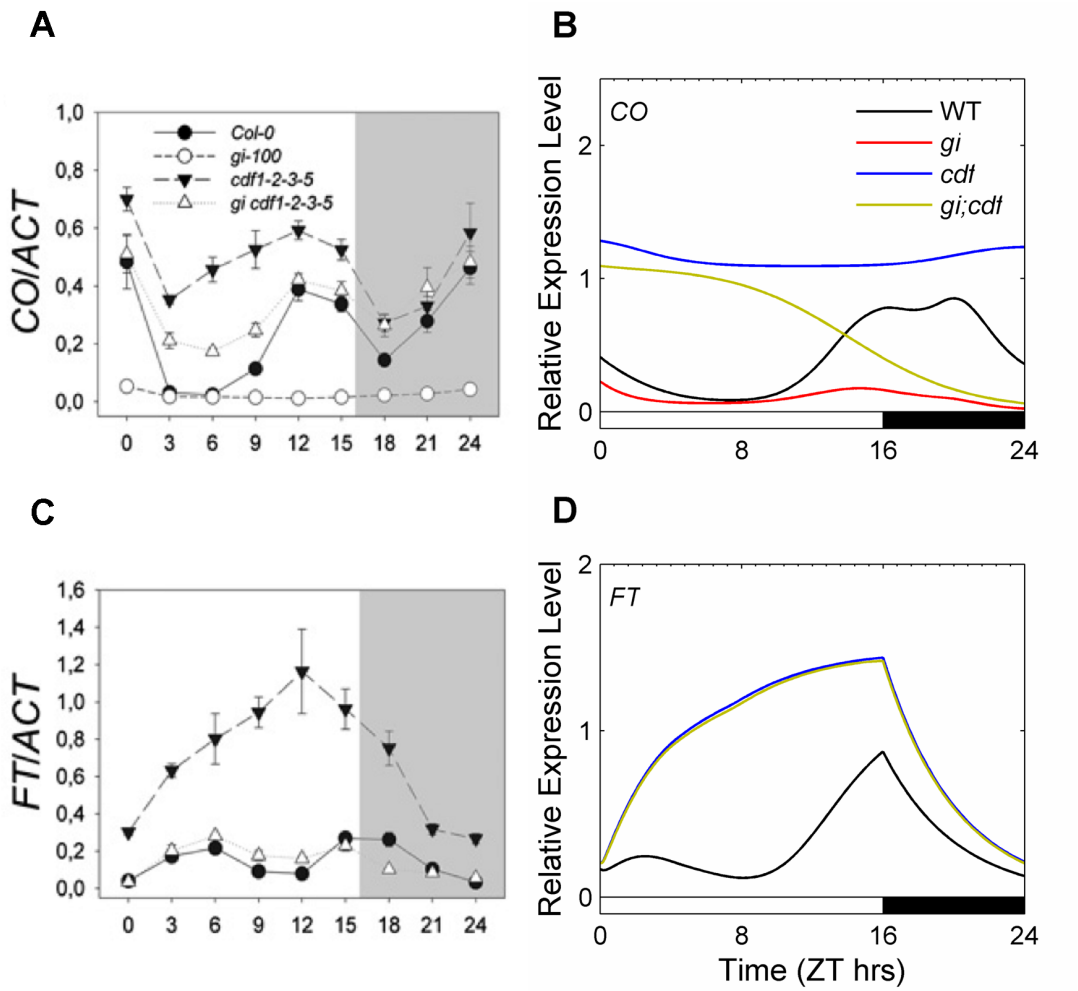


Figure 3.4: Simulations of *gi* and *cdf* mutants. (A) Data for *CO* transcription in LDs in WT (solid black line, filled circles), *gi* (dashed black line, empty circles), *cdf1-R;2;3;5* (dashed black line, filled triangles) and *gi;cdf1-R;2;3;5* (solid grey line, empty triangles). (B) LD simulations of *CO* mRNA in WT (black line), *gi* (red line), *cdf* (used as a proxy for *cdf1-R;2;3;5*, blue line) and *gi;cdf* (yellow line). (C and D) as in (A) and (B) for *FT* mRNA. (A and C) taken from [83]. Errorbars represent standard deviation.

- *Revised regulation of CO protein and FT transcription*

The updated flowering time model equations are:

$$\frac{dc_{CDF1}}{dt} = p_1 c_{CDF1}^{(m,D)} - p_2 L c_{FKF1}^{(D)} c_{GI} c_{CDF1} - m_1 c_{CDF1} \tag{3.1}$$

$$\frac{dc_{CO}^{(m)}}{dt} = B_{CO} + n_1 \frac{g_1^a}{g_1^a + c_{CDF1}^a} \frac{c_{GI}}{g_2 + c_{GI}} + n_2 \frac{c_{TOC1}^b}{g_3^b + c_{TOC1}^b} - m_2 c_{CO}^{(m)} \tag{3.2}$$

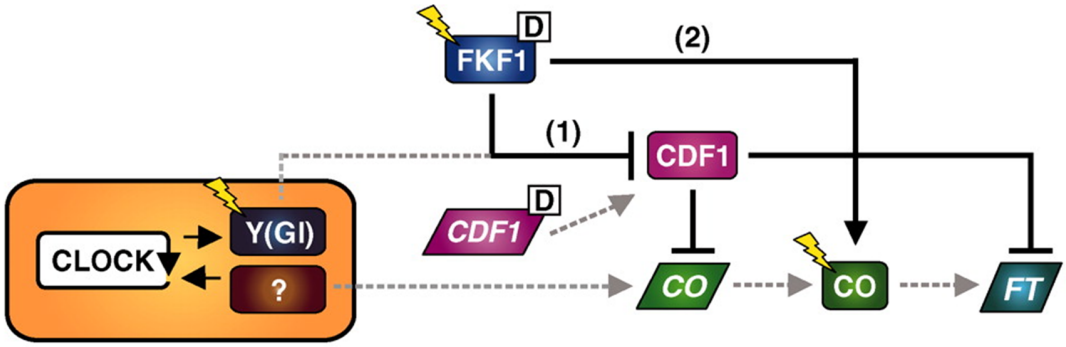


Figure 3.5: Schematic of photoperiodic flowering model. The 'Locke2005' circadian clock model (in orange) is connected to the flowering pathway through a GI-FKF1 protein interaction and activation of *CO* transcription through an unknown mechanism (highlighted by a "?" in the schematic, *TOC1* mRNA in the 'Locke 2005' model). Parallelograms represent mRNA components; rectangles represent protein components. Thunderbolts represent light-regulated components. The two roles of FKF1 in the system are highlighted: (1) FKF1-CDF1 interaction; (2) FKF1-CO interaction. Components marked with a "D", *CDF1* mRNA and FKF1 protein, highlight the data inputs into the model.

$$\frac{dc_{CO}}{dt} = p_3 c_{CO}^{(m)} - p_4\left(m_3 + m_4 D - L\frac{c_{FKF1}^{(D)}}{g_4 + c_{FKF1}^{(D)}}\right)c_{CO} \tag{3.3}$$

$$\frac{dc_{FT}^{(m)}}{dt} = \left(n_3 + n_4\frac{g_5}{g_5 + c_{CDF1}}\right)\frac{c_{CO}^{c}}{g_6^{c} + c_{CO}^{c}} - m_5 c_{FT}^{(m)} \tag{3.4}$$

where n_i and p_i represent transcription and translational rates respectively; m_i are the degradation rates; g_i are the Michaelis-Menten constants interpreted as binding affinities; a, b and c are the Hill coefficients; B_{CO} is the basal transcription rate of *CO* mRNA [82]; L and D represent light and dark from the light function present in the model of the circadian clock, connecting the system to the photoperiod. Superscript (D) signifies components of the model that are read into the system directly from experimental data and superscript (m) represents those components that are mRNA.

The data presented in Figure 3.6 highlight that: the CO protein is stabilized before dusk by FKF1, altering the waveform of *FT* mRNA (Figure 3.6a-c), and; CDF1 is active at a similar region of the *FT* promoter as CO and FKF1 at the same time

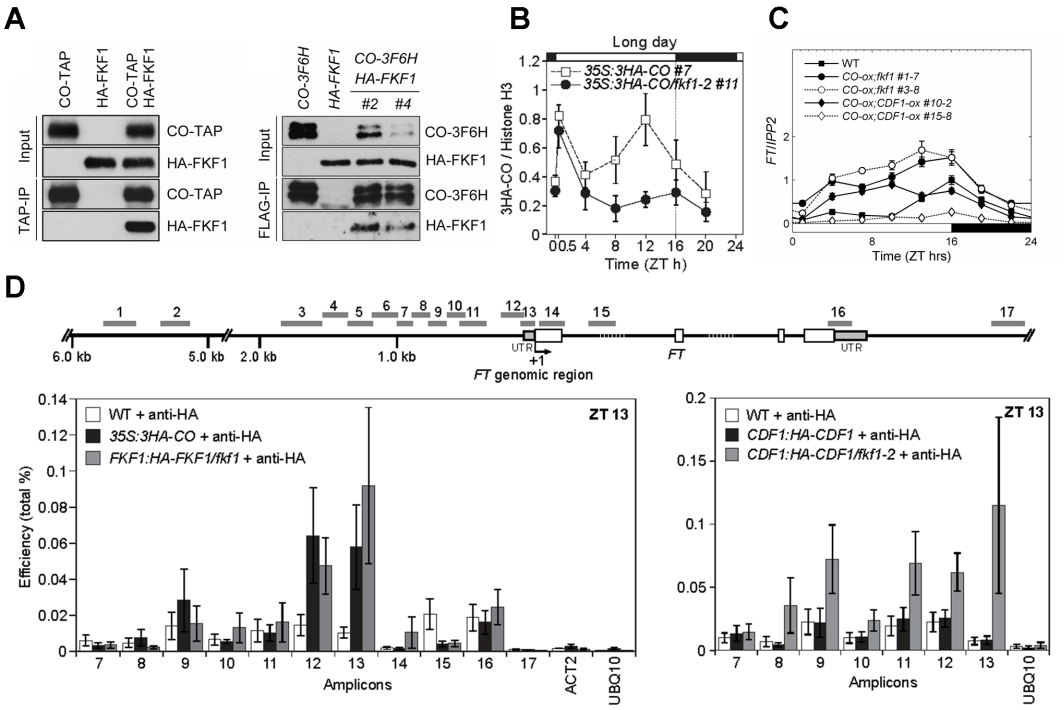


Figure 3.6: CO, CDF1 and FKF1 all play roles in the regulation of *FT* mRNA. (A) Immunoprecipitation experiments showing that CO and FKF1 proteins are able to interact. (B) CO protein is destabilised at the end of long days when FKF1 is absent. *CO-ox* = dashed black line, empty squares; *CO-ox;fkf1* = solid black line, filled squares. (C) *CDF1-ox* suppresses rhythms of *FT* transcription in *CO-ox* lines more strongly than the removal of FKF1 function. WT = solid black line, filled squares; *CO-ox;fkf1* = circles; *CO-ox;CDF1-ox* = diamonds. (D) ChIP experiment showing that CO, FKF1 and CDF1 proteins are all active at the same locations on the promoter of *FT* at ZT13 of LD conditions. Activity of CDF1 is increased in the absence of FKF1 function. Error bars represent standard error.

suggesting that these factors may act antagonistically to each other to regulate *FT* expression (Figure 3.6d). The antagonism between CDF1 and CO functions can be observed in Figure 3.6c, where the suppressive effect of constitutively overexpressed *CDF1* mRNA is seen even when *CO* mRNA is also overexpressed. These observations supports the choice of a multiplicative interaction among these regulators to control *FT* transcription in the model as *FT* mRNA depends almost completely on CO protein, and is modelled to ensure that there is no *FT* expression produced in the *co* mutant background (equation 3.4) [73]. This result also shows that transcriptional activation of *FT* is sensitive to inhibition by CDF1. Hence, according to the model proposed here, FKF1 has a dual role in regulating *FT* mRNA: 1. degrading CDF1 protein to alleviate inhibition of *CO* and *FT* mRNA (equations

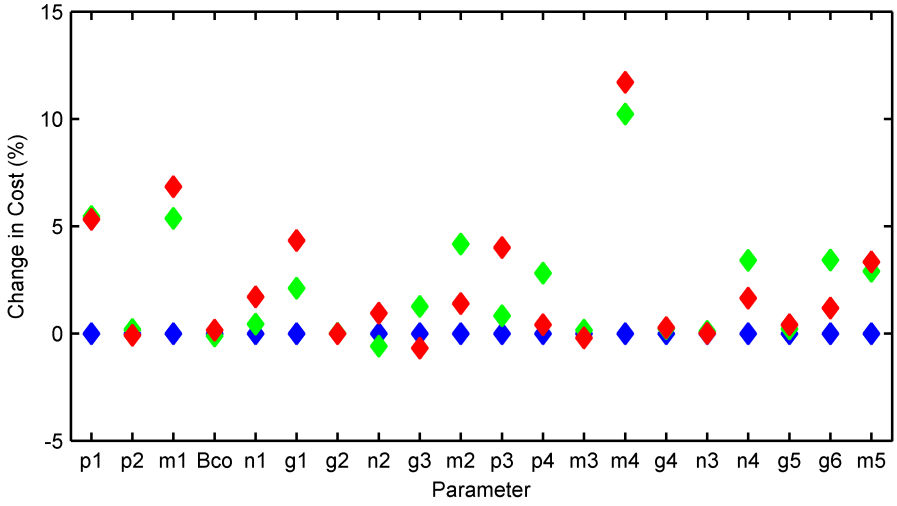


Figure 3.7: Sensitivity of *FT* mRNA rhythms to perturbations of single parameter values. The percentage change in the goodness-of-fit (cost) of *FT* mRNA simulations compared to the training data was recorded after altering the model parameter values with a 10% increase (green diamonds) or decrease (red diamonds) compared to their WT values (blue diamonds).

3.1 and 3.2), and; 2. stabilizing CO protein to regulate *FT* mRNA (equation 3.4), as depicted schematically in Figure 3.5. During darkness, the CO protein is strongly degraded by a COP1 dependent mechanism [87, 88]. Therefore, the phase of FKF1 has a relationship with the phases of the external environment (e.g. light) and also determines the phases of *CO* and *FT* mRNA that are internally coincident to FKF1 [207].

Computational methods

As with many other mathematical models of gene regulatory systems, the full set of kinetic parameters has not been measured through experimentation. Parameters can be approximated or constrained from the data sets, within biochemically reasonable bounds. Due to the consistency seen among the datasets used in this study (seen in Figure 3.1) from varying genotypes and photoperiods, these provide the strongest constraints on parameter values in the model. Although the resulting parameter values are not uniquely specified, they are able to accurately describe the dynamics of the biological system enabling the model to be analyzed and make predictions

about the flowering pathway. I will show below that our predictions and analysis still hold even when parameter values are varied.

The *CO* and *FT* mRNA equations were fitted to the same published data sets that were used in the optimization of the original flowering time model [82]. In order to determine the relationship between the parameters n_1 and n_2, representing the two rhythmic components of *CO* transcription, *CO* mRNA equation was fitted to data from both WT plants and *fkf1* mutants. This ensured that the CDF1 term did not dominate the *CO* mRNA equation. In a similar manner, to obtain the correct relationship between p_2 and m_1, the CDF1 protein equation was fitted to data from WT and *fkf1*. The three Hill coefficients were set to 2, representing dimeric binding of the transcriptional regulators [12]. Another constraint that was placed on the system was that $g_4 = 3g_1$ since the binding efficiency of CDF1 protein to the promoter of *FT* appears to be ~33% as efficient as binding to the *CO* promoter (compare Figure 3.6d to results in [81]), though other ratio's could be chosen without greatly affecting the model. Due to low abundance, data for CO protein in the WT background is lacking, hence parameter values for equation (3.3) were obtained by fitting equation (3.4) to *FT* mRNA data. In total, this meant that 18 parameters were fitted to 96 data points from the 12 WT and *fkf1* data sets taken from [82]. To validate the model, the remaining 44 data sets (consisting of 352 data points) were used. The resulting parameter values are given in Appendix A.

The models were optimized using the simulated annealing algorithm simulannealbnd from the Matlab R2008b Optimization Toolbox (Mathworks, Cambridge, UK). Simulated annealing is able to find local optima for the models parameter set, providing an accurate fit for the model compared to data (see Chapter 2; [208]). For this study I have used the slower Boltzman annealing procedure with an exponential 'temperature' update starting at an initial 'temperature' of 1. Goodness of fit was calculated by a Euclidean difference/norm that I refer to as the "cost" of the parameter set, hence a lower cost gives a better-fitting model (see Chapter 2). When each of the obtained parameter values was increased (green diamonds) or decreased (red diamonds) by 10%, the changes caused moderate increases in the cost of *FT*

simulations under LD conditions and no significant decreases in goodness-of-fit, illustrating the robustness of the model to individual parameter changes (Figure 3.7).

Results

Experimental evidence of feed-forward network

As discussed in the Methods section above, collaborators in Assistant Prof. Takato Imaizumi's group (University of Washington, USA) provided numerous datasets to aid construction of the model described in this chapter. These have been published in [153], however the key findings will be summarised here. First, CO and FKF1 proteins are able to form a complex that stabilises CO protein levels at the end of long days (Figure 3.6a and b). This occurs mainly through the LOV domain of FKF1 [153]. Furthermore, association of FKF1 was recorded on the *FT* promoter in the same regions as where CO binds and regulates *FT* transcription at ZT13 in LDs (Figure 3.6d). Thus, this validates the model derived prediction made by Salazar and co-workers that FKF1 helps promote the CO-dependent activation of *FT* transcription [82].

Second, the high expression of *FT* mRNA in *CO-ox* backgrounds can be suppressed by removing FKF1 activity or by overexpressing *CDF1* transcription (Figure 3.6c). The strong inhibition of *FT* transcription due to *CDF1* overexpression suggests that CDF1 protein suppresses *FT* mRNA levels in a similar manner to *CO* transcription. This mechanism is proposed to occur on the promoter of *FT*, where CDF1 associates with the promoter at the same location and time as CO and FKF1 proteins (Figure 3.6d). The *fkf1* mutant has lower levels of CO and higher levels of CDF1 protein, respectively, prior to dusk (Figure 3.6b and 3.8b). This results in reduced *FT* expression at dusk due to the lower abundance of a positive regulator, CO, and higher levels of a negative regulator, CDF1 (Figure 3.6c). Thus, both CDF1 and FKF1 form coherent feed-forward loops regulating both *CO* and *FT* transcription (Figure 3.5).

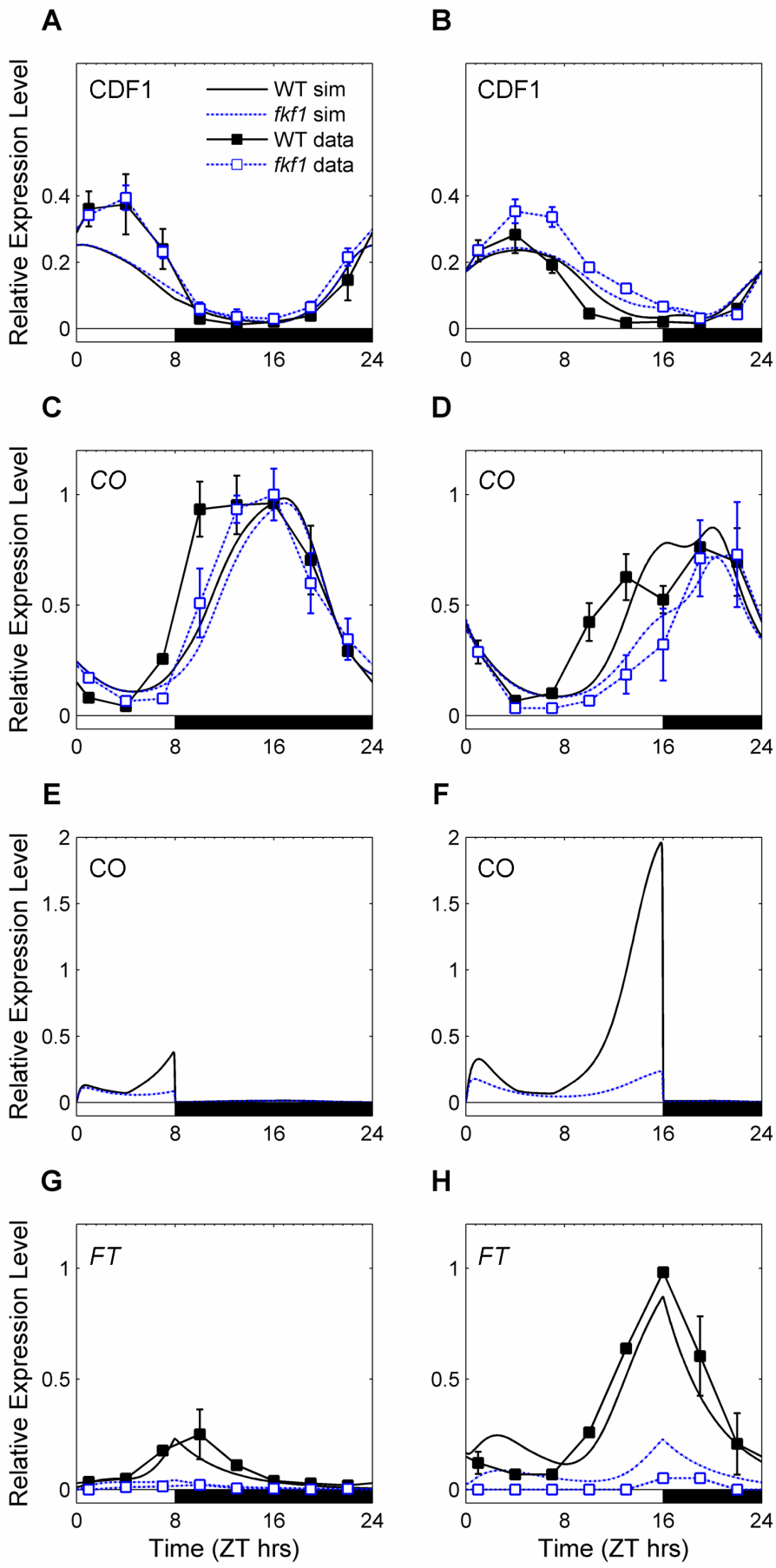
Comparison to the previous photoperiod response model

In comparison to the previous flowering time model (termed 3F2 in [82]), the new mathematical model significantly improves the biological realism of the flowering time system by modelling the biochemical mechanisms of FKF1 and CDF1, which were absent from the earlier model. As seen in Figure 3.8, the simulations from the model developed here for WT CDF1 protein, *CO* mRNA, CO protein and *FT* mRNA qualitatively match the training datasets. Importantly, simulations of *CO* mRNA describe the FKF1-dependent 'shoulder' at ~ZT13 of LDs (see above) that was captured in the previous model, albeit with a slight delay [82].

Since the previous model was also able to qualitatively describe the *fkf1* mutant, it was important that the model developed here retained the ability to match these datasets. As seen in Figure 3.8, simulating *fkf1* mutants led to the loss of the first *CO* peak at ZT13 in LD conditions in a similar manner to the previous model. The expression of *FT* mRNA was also correctly suppressed in *fkf1* mutants due to the FKF1 feed-forward loop. Thus, the new model detailed here was as good as the previous system in describing WT and *fkf1* backgrounds [82]. With the extra detail that has now been shown for the flowering time system, I wished to be able to describe further genetic perturbations to the network.

Describing rhythms resulting from genetic perturbations

As well as data collected from WT and *fkf1* loss-of-function plants, new data from transgenic plants overexpressing *CO* transcription (*CO-ox*) crossed with *fkf1* mutants (*CO-ox;fkf1*) and a constitutive *CDF1* overexpressor (*CO-ox;CDF1-ox*) allowed us to further constrain the model and the roles of FKF1 and CDF1 proteins in regulating *FT* mRNA (see Figure 3.6c and Figure 3.9). The previous model had not been built to describe these datasets and did not feature *CDF1* mRNA or protein as components of the system (see above). In order to describe *CO* overexpression, the basal transcription rate (parameter B_{CO}) was increased to 150 times its normal value. This meant that the rhythm of CO protein seen in overexpression lines would be


A
B
C
D
E
F
G
H
CDF1
CO
FT
WT sim
fkf1 sim
WT data
fkf1 data
Relative Expression Level
Time (ZT hrs)
0
8
16
24
0.2
0.4
0.5
1
1.5
2

maintained with an elevated mean expression level (Figure 3.6b and Figure 3.9c & d; [86]). As seen in Figure 3.2c, the level of *CDF1* mRNA in the *CO-ox*;*CDF1-ox* background is greatly increased compared to other backgrounds. Hence, to describe the *CO-ox*;*CDF1-ox* double overexpression line, the input $c_{CDF1}^{(m,D)}$ was changed from the WT data to data from the *35S:CDF1;SUC2:CO* line. The *CO-ox*;*fkf1* can then be simulated by setting $c_{FKF1}^{(D)} = 0$ in equations (3.1) and (3.3). From Figure 3.9, the model simulations of *CO* and *FT* mRNA in all these backgrounds are qualitatively similar to published experimental data (Figure 3.6; [153]). Figure 3.9f shows that simulating *CO-ox* leads to rhythms of *FT* transcription with higher mean levels that still gain a marked end-of-day peak in LDs, matching the data of Figure 3.6c. The model is also able to match the rhythms of CO protein observed in *CO-ox* and *CO-ox;fkf1* mutants (compare Figure 3.9d with Figure 3.6b). Similarly, the model qualitatively describes the LD *FT* transcription profile in *35S:CO;fkf1* whereby expression simply increases in the light and falls in darkness (Figure 3.9f; compare to Figure 3.6c), reflecting the strong light regulation of CO protein. Importantly for our analysis, Figure 3.8f also shows that the effect of *CDF1-ox* on *CO-ox* simulations was greater than the effect of *fkf1*, matching the data seen in Figure 3.6c.

Model predicts effects of FKF1 overexpression

A good way to test how accurate a model is at describing a biological system is to experimentally validate a hypothesis generated from the model. Here we tested whether the model could predict the impact of constitutively overexpressing *FKF1* transcription on the rhythmic expression of *FT* mRNA. The model simulated two

---

(continued from above) Figure 3.8: Simulations of model components compared to training data in WT and *fkf1* loss-of-function mutants. Simulations of (A, B) CDF1 protein; (C, D) *CO* mRNA; (E, F) CO protein, and; (G, H) *FT* mRNA in (A, C, E, G) SD and (B, D, F, H) LD conditions from WT and *fkf1* conditions were compared against the available training data. WT data = solid black line, filled squares; *fkf1* data = dashed blue line, empty squares; WT simulations = solid black line; *fkf1* simulations = dashed blue line. Data for CDF1 protein replotted from Figure 3.2d, data for *CO* and *FT* mRNA taken from Figure 3.1 [74]. Error bars represent standard error.

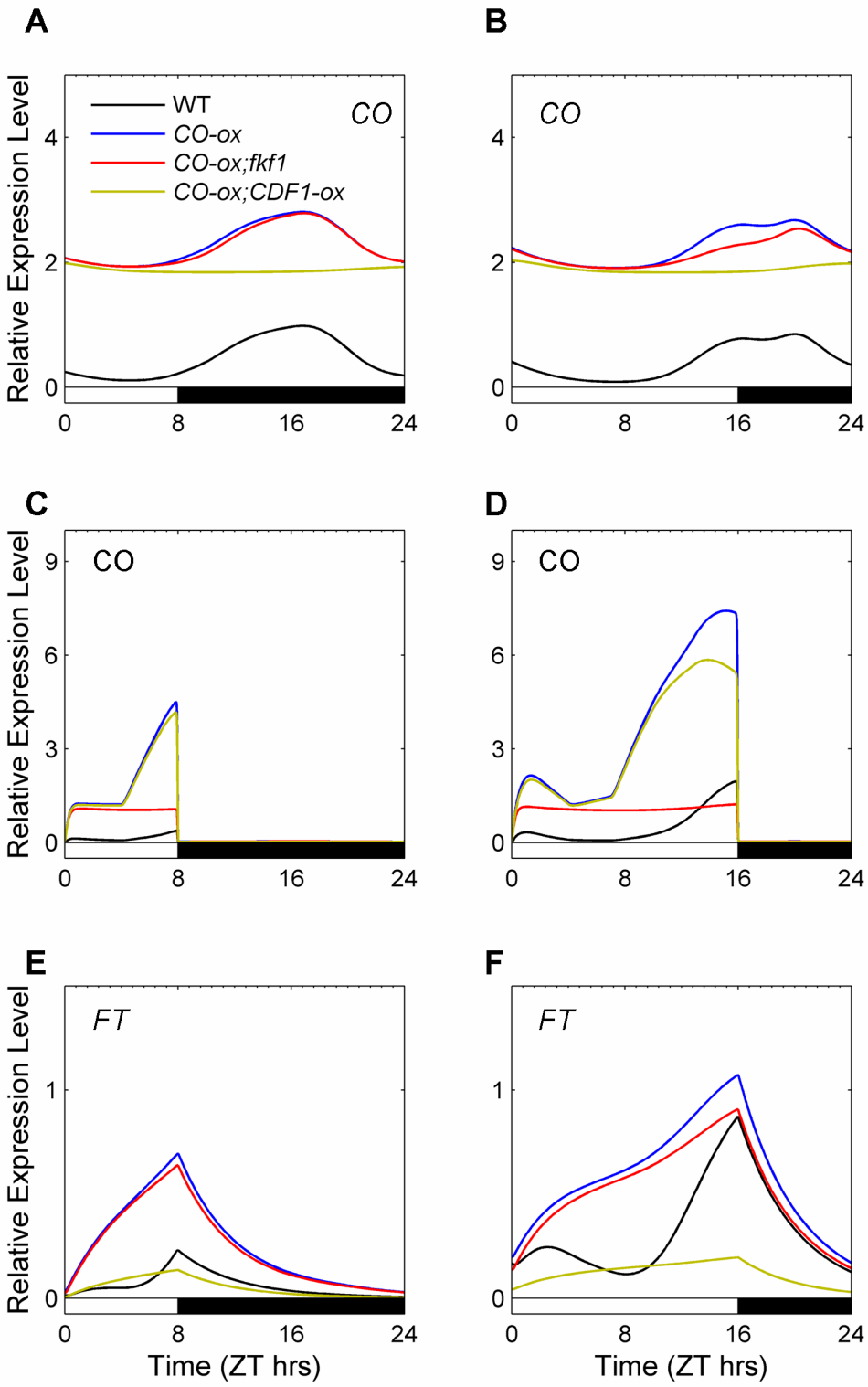


Figure 3.9: Simulations of model components in *CO* overexpression transgenic plants. (A, B) *CO* mRNA, (C, D) CO protein and (E, F) *FT* mRNA in (A, C, E) SD

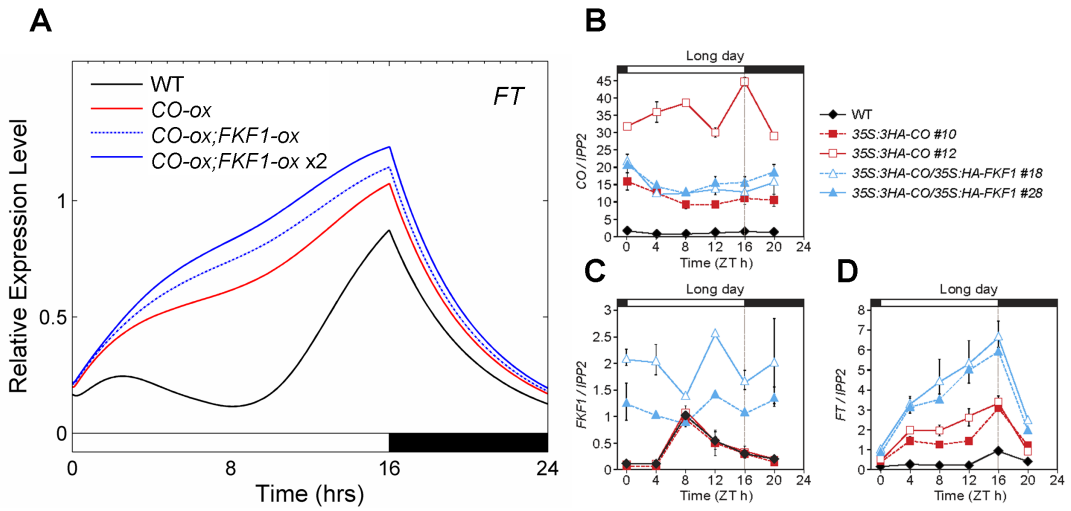


Figure 3.10: Model predicted rhythms of *FT* mRNA in FKF1 overexpression lines validated experimentally. (A) Simulations of LD *FT* mRNA rhythms in WT (black line), *CO-ox* (red), *CO-ox;FKF1-ox* (dashed blue) and *CO-ox;FKF1-ox x2* (solid blue) that has double the level of *FKF1* mRNA. (B-D) Levels of (B) *CO* mRNA, (C) *FKF1* mRNA and (D) *FT* mRNA determined experimentally in each of the tested lines – WT = black line, filled diamond; *CO-ox* = red line, squares; *CO-ox;FKF1-ox* = blue lines, triangles. Error bars represent standard error.

FKF1-ox lines in a *CO-ox* background by setting the input value $c_{FKF1}^{(D)}$ to the maximum value (or twice the maximum value) of WT FKF1 protein levels throughout the full diurnal cycle (Figure 3.10a). This would also remove the circadian regulation of FKF1 that is seen in WT data and highlight that circadian regulation of FKF1 was important for the correct timing of *FT* mRNA. As can be seen in Figure 3.10a, increasing the levels of *FKF1-ox* had a positive effect on *FT* expression during the day due to the increased stability of CO protein. Therefore, removing the circadian regulation of FKF1 leads to more *FT* being produced early in the photoperiod. This hypothesis was tested by Assistant Prof. Takato Imaizumi's group (University of Washington, USA), who showed that the effect of *35S:HA-FKF1* in *35S:3HA-CO* transgenic plants was strongest on *FT* mRNA rhythms during the day of LD cycles and was dependent on the level of *FKF1* in each of the lines (Figure 3.10c & d). Since the amount of *CO* mRNA present in the two *35S:3HA-CO;*

(continued from above) and (B, D, F) LD conditions simulated in WT (black lines), *CO-ox* (blue), *CO-ox;fkf1* (red) and *CO-ox;CDF1-ox* (yellow) backgrounds. Error bars represent standard error.

35S:HA-FKF1 lines was comparable then the effects on *FT* expression are most likely driven by FKF1 stabilization of CO protein levels or activity (Figure 3.10b). The validation of a hypothesis derived from the model, in addition to the models ability to describe numerous expression profiles from transgenic mutants, illustrates that our model provides a qualitative representation of the biological system that underlies photoperiodic flowering. Hence, using the model, I wished to compare the importance of the two roles for FKF1 protein in the regulation of *FT*.

Estimating the importance of FKF1 in the model

To provide a quantitative measure of the effects of FKF1 on the system, I calculated the changes in the area under the waveform of *FT* mRNA over a single LD cycle (termed FT_{AREA}). In the model, FKF1 has two roles: 1. to degrade CDF1 protein, and; 2. to stabilize CO protein. Hence, two partial mutants were created that removed one of these two processes whilst maintaining the other (denoted Δ(1) and Δ(2), respectively). To perform such an analysis experimentally would be difficult and require a lot of information about the protein structure of FKF1. By simulating Δ(1), Figure 3.11a shows that FT_{AREA} decreased by ~22% in comparison to simulated WT levels of *FT* mRNA. However, by simulating the Δ(2) mutant, there was ~52% decrease in FT_{AREA} and ~48% less *FT* mRNA at dusk (Figure 3.11a). Hence, the stabilization of CO protein by FKF1 is relatively more important than the degradation of CDF1 by FKF1 for the correct regulation of *FT* transcription. However, as seen by the percentage decreases of FT_{AREA}, the degradation of CDF1 by FKF1 still plays a significant role in the regulation of *FT* mRNA (Figure 3.11a).

As discussed in the Methods section above, variation in parameter values can lead to variations in model results. To confirm conclusions about the two roles of FKF1, I calculated FT_{AREA} in the partial mutants for 500 quasi-random parameter sets selected from a Sobol series (Figure 3.11b). Parameter sets were scored for the cost of their simulations to the training data (*CO* and *FT* mRNA profiles, and CDF1 protein rhythms), and those that also predicted a larger FT_{AREA} in *CO-ox/fkf1* than

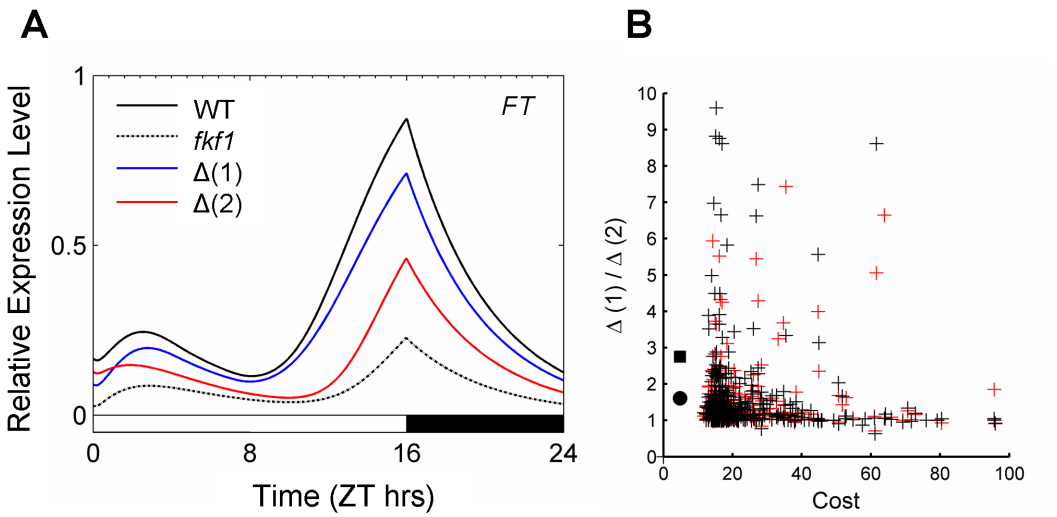


Figure 3.11: Simulations of partial *fkf1* loss-of-function mutants separates the relative contributions of the two FKF1 mechanisms that regulate *FT* mRNA. (A) Simulations of LD *FT* mRNA in WT (solid black line), *fkf1* (dashed black line), Δ(1) (blue) and Δ(2) (red) that are partial *fkf1* mutants. Δ(1) = removal of FKF1-CDF1 interaction; Δ(2) = removal of FKF1-CO interaction. (B) Ratio of LD FT_{AREA} (black crosses) and peak *FT* expression level (red crosses) simulated in the two partial mutants was calculated for 500 quasi-random parameter sets. Ratio of LD FT_{AREA} (square) and peak *FT* expression (circle) from (A) are highlighted.

CO-ox/CDF1-ox, as had been observed in Figures 3.6c and 3.9f, were retained. Figure 3.11b shows that the majority of retained parameter sets produced simulations with a smaller FT_{AREA} in Δ(2) than Δ(1), with a lower amount of *FT* mRNA being expressed at dusk, as had been the case with the initial parameter set (represented by the square and circle in Figure 3.11b, respectively). The notable exceptions to this conclusion occurred when the cost was significantly high such that the resulting simulations would not prove to be a good match to the training datasets. Thus the CO stabilization mechanism has a significant effect in all parameter sets that accurately describe the training data.

As removal of the CO-FKF1 interaction has a large effect on *FT* mRNA expression levels in LDs, I wished to determine whether constitutively stabilized CO protein (by FKF1) would result in *FT* transcription levels that are less sensitive to photoperiod. To do this, I simulated a constant CO-FKF1 interaction by setting c_{FKF1} in equation (3.3) to be maximally constant throughout the diurnal cycle and calculated the

differences in FT_{AREA} between SD and LD conditions (FT_{AREA}^{SD} and FT_{AREA}^{LD} respectively). The resulting simulations showed that in WT conditions FT_{AREA}^{SD} is ~23% of FT_{AREA}^{LD}, highlighting the large difference in rates to flowering of Arabidopsis in the two day-lengths [72]. However, when the stabilization of CO occurs throughout the whole day, FT_{AREA}^{SD} increases to be ~46% of FT_{AREA}^{LD} (and ~63% of the WT FT_{AREA}^{LD}). This implies that the circadian regulation of FKF1 protein is important for photoperiodic flowering, correctly activating *FT* expression in LD conditions and preventing premature floral induction in SDs.

Discussion

Updated flowering time model is able to differentiate between redundant roles of FKF1

In this chapter, a mathematical model of the photoperiod-dependent flowering time system in *Arabidopsis thaliana* has been built to quantify the roles of different proteins in the system. The previous model predicted the blue light-dependent FKF1 E3 ligase plays a dual role in the correct timing of floral initiation. First, FKF1 is crucial for the activation of *CO* mRNA prior to dusk in LD conditions by degrading the transcriptional inhibitor CDF1. Second, FKF1 forms a light-dependent protein complex with CO that stabilizes CO on the promoter of *FT* whose mRNA goes on to regulate flowering. The Imaizumi group (University of Washington, USA) have experimentally proven this hypothesis and, furthermore, shown that CDF1 also plays a secondary role in the flowering system by inhibiting *FT* in a CO-dependent manner (Figure 3.6). The dual roles of FKF1 and CDF1 thus form feed-forward networks in the extended flowering system [89].

The new mathematical model was built to include regulation of CDF1 and CO protein levels by FKF1 in the flowering system. Simulations of the model in comparison to data from WT and transgenic plants show good qualitative matches (compare Figures 3.6c and 3.9f and Figure 3.8). The model also predicted that

overexpression of FKF1 would lead to increased levels of *FT* mRNA specifically during the day. This hypothesis was validated experimentally (Figure 3.10). Furthermore, by using the area under the curve of *FT* mRNA (FT_{AREA}) from our simulations we were able to quantify the effects of altering FKF1 in the system. For example, the removal of the circadian regulation of FKF1 protein means that there is relatively more *FT* mRNA in SD compared to LD conditions than there is when FKF1 protein maintains rhythmic expression levels. Hence, the circadian timing of FKF1 is critical for the suppression of floral induction in SDs that mimic day-lengths observed in winter months when floral induction would not be advantageous for *Arabidopsis thaliana*.

In a similar manner, by using *FT* mRNA as a readout for the effects of FKF1 misexpression, we were able to show that the stabilization of CO protein by FKF1 had a relatively larger effect on *FT* levels than CDF1 degradation by FKF1. However, removal of FKF1-dependent degradation of CDF1 still had a noticeable effect on *FT* expression suggesting that it is still important in the flowering system. This implies that the multiple and partially redundant roles of FKF1 lead to a strong switch in floral initiation as growth conditions change seasonally from SDs to LDs.

Limitations of the model

The model presented in this chapter was built with the specific purpose of trying to quantify the effects of FKF1 on the flowering time pathway. As a consequence, other than adding extra detail to the previously published model, several limitations have not been addressed. An example of this is the regulator of the LD specific night-time peak of *CO* mRNA. In both the previous model and the model presented in this chapter, this was created artificially using a component of the circadian clock model that tracks dusk in the correct manner necessary to generate this peak. There is currently little in the way of published information as to what the regulator of this second peak could be, however *CO* mRNA in *cop1* and *fbh* mutants lose only this second peak in long days implying that COP1 and/or FBH proteins may play a role in this regulation by activating *CO* transcription at night [34, 85].

A further limitation is the lack of quantified data on the low-abundance CO protein in WT plants under a range of light conditions. Previous studies have shown that CO protein is regulated by phytochrome and cryptochrome photoreceptors whilst also being regulated by COP1 and the FKF1 stabilization mechanism presented here [86–88, 153]. In addition to its peak at the end of the day, CO protein levels can also peak acutely after the transition from dark to light. This is very marked in the *35S:CO* lines (Figure 3.6b). It is simulated to a lesser extent in the model, even in WT conditions, owing to the significant level of *CO* mRNA still present at dawn in LDs (Figure 3.8f). Several studies have shown significant *FT* activation around ZT4 in LD grown WT plants, presumably due to further regulation of CO protein [75, 82, 86].

Whilst the model presented here is more flexible and descriptive than the previous system, the main limitation in the models flexibility is that two components are not mathematically modelled (components marked "D" in Figure 3.5). Data sets are required as inputs to the system for *CDF1* mRNA and FKF1 protein as the way they are regulated by the circadian clock has currently not been fully elucidated. This lack of a fully mathematical model limits the number of photoperiods and circadian clock mutants that can be simulated. The model presented in this chapter is able to describe *CO* mRNA in a *gi* loss-of-function mutant and conditions in 8hr photoperiods (SD) and 16hr photoperiods (LD) but no other clock mutants or photoperiods. As was discussed in Chapter 2, to be able to more accurately describe variations in clock architecture via genetic perturbations and varying photoperiods a larger clock model should be used in the model of output pathways. Thus, one of the aims of Chapter 4 is to build on the model presented here. By building models for the circadian regulation of *CDF1* and *FKF1* and updating the circadian clock model used in the system more photoperiods and circadian clock mutants will be able to be described.

Chapter 4: Mathematical Modelling of the Thermo-Photoperiodic-dependent Flowering System in *Arabidopsis thaliana*

This work has been submitted in the following manuscript:
R. W. Smith, D. D. Seaton, Y. H. Song, D. R. MacGregor, K. Stewart, G. Steel, J. Foreman, S. Penfield, T. Imaizumi, A. J. Millar, K. J. Halliday. Linked circadian outputs control growth and development in response to photoperiod and temperature.

In the previous chapter, a mathematical model was introduced and discussed that focussed on the circadian- and light-regulation of the molecular pathway controlling photoperiodic flowering in Arabidopsis (see Chapter 3; [153]). This model was able to correctly describe the flowering pathway in a number of flowering-specific mutants and in two different photoperiods. Furthermore, the model correctly predicted rhythms of *FT* mRNA in constitutive *FKF1* and *CO* overexpression lines. However, as the model required data sets for *CDF1* mRNA and FKF1 protein to be input into the model to work, this system would be unable to simulate conditions for which data was unavailable. This limits the usefulness of the model. Here, we build on this model to include the circadian regulation of all components in the system, removing the need for data inputs, and then test how warm ambient temperature regulates flowering in experimental conditions that mimic summer months.

Circadian regulation of CDF1 and FKF1 transcription

A number of published data sets show that *CDF1* and *FKF1* mRNA are regulated by the circadian clock. Data from WT plants, as well as *cca1;lhy* and multiple *prr* loss-of-function mutants in both 10L:14D and 16L:8D diurnal cycles were available (throughout this chapter both 10L:14D and 8L:16D shall be referred to as SDs; [50, 78, 209]). In *cca1;lhy* mutants, both *CDF1* and *FKF1* transcription has an advanced

phase such that peak expression occurs ~3hrs earlier than in WT rhythms. Conversely, in the *prr9;7* loss-of-function transgenic plants, both *CDF1* and *FKF1* transcription rhythms have a delayed phase and a broader peak of expression. In combination with ChIP experiments, the published literature supports a mechanism whereby *CDF1* transcription is repressed by the PRR proteins [30, 31]. The regulation of *FKF1* mRNA by the circadian clock, on the other hand, closely correlates with the regulation of *GI* transcription such that the peak of expression occurs at the same time of the day, both *GI* and *FKF1* respond in the same manner to circadian clock perturbations and they are both acutely activated by red light [23, 78, 80, 142, 209]. In the results section of this chapter I will model these modes of regulation within the flowering model.

Warm temperature-controlled acceleration of flowering through FT levels

As discussed in Chapter 1, PIF4, and to a lesser extent PIF5, have been implicated in promoting hypocotyl elongation, the seedling stem (see Chapter 2), in response to warmer ambient temperatures (~27°C; [122, 123, 161, 174]). The rise in temperature correlates with increasing *PIF4* mRNA & protein levels and, consequently, longer hypocotyls [122, 123, 125, 161, 162]. Recently, PIF4 has also been implicated in the warm temperature-induced acceleration of flowering in SD conditions [174]. The elevation in temperature leads to accelerated flowering due to increased *FT* transcription in a PIF4-dependent manner. Interestingly, the PIF4 regulation of *FT* has been postulated to be independent of the photoperiodic flowering pathway as *co* and *gi* loss-of-function transgenic plants maintain temperature-sensitive flowering phenotypes in SD [173, 174]. However, in LD conditions, the *co;PIF4-ox* transgenic plant shows partially additive phenotypes such that the removal of CO delays the accelerated flowering observed by PIF4 overexpression [174]. This suggests that PIF4 may act, in part, with the photoperiod-dependent pathway to regulate flowering in Arabidopsis.

In this chapter the flowering model discussed in Chapter 3 will be expanded to: 1. include the circadian regulation of *CDF1* and *FKF1* transcription, and; 2. analyse the

mechanism through which PIF4 regulates *FT* and temperature-accelerated flowering in LD conditions when CO regulation of *FT* is crucial for flowering to take place. After building this detailed model, I will probe the system and develop hypotheses that can be tested experimentally.

Methods

Experimental Methods

New transcript data for *CO* and *FT* mRNA was obtained by Dr. Kelly Stewart, Mr. Gavin Steel and Dr. Julia Foreman from Dr. Karen Halliday's group (University of Edinburgh, UK). Chromatin immunoprecipitation assays were conducted by Dr. Dana MacGregor, a collaborator on the ROBuST project from Dr. Steve Penfield's group (University of Exeter, UK). New CDF1 protein data acquired for this study was obtained by international collaborators Dr. Young Hun Song and Assistant Prof. Takato Imaizumi (University of Washington, USA).

- *Growth Conditions for RNA analysis*

Seeds of WT (Col-4, Columbia accession) and *pif4;5* plants were surface sterilised then 30-40 seedlings were sown on 55 mm diameter plates containing half-strength MS media (Melford, Ipswich, UK) , pH 5.8, and 1.2% agar without added sucrose. The seeds were stratified at 4°C for 3 days and then grown for 13 days in 16L:8D cycles (100 µmol m^{-2} s^{-1} from cool white fluorescent tubes) at 22°C and 27°C. Seedlings were harvested from triplicate samples at ZT0, 1, 2, 4, 8, 12, 15, 16, 17, 20 and 24 (ZT = zeitgeber time, ZT0 = lights on) into RNA*later* solution (Sigma-Aldrich, Gillingham, UK). The plants were left overnight at 4°C in the RNA*later* solution to allow full penetration into the tissue [23]. The generation and growth of *35S:3HA-CO* constructs has been previously described [153].

- *RNA Extraction*

RNA was extracted from the plant tissue using the Illustra RNAspin 96 RNA isolation kit (GE Healthcare, Chalfont St. Giles, UK) manually, as described [210]. Purified total RNA (1 μg) was reverse transcribed into cDNA using SuperScript VILO cDNA synthesis kit with oligo dT primers (Invitrogen/Life Technologies, Paisley, UK) according to the manufacturer's instructions. cDNA was diluted 1/10 and 1 μl used for subsequent qRT-PCR.

- *Gene expression analysis*

qPCR reactions were set up using a liquid handling robot (freedom Evo, TECAN, Reading, UK) and run in a Lightcycler 480 system (Roche, Burgess Hill, UK) using LightCycler 480 SYBR green master mix (Roche, Burgess Hill, UK). Data was analysed with Roche Lightcycler 480SW 1.5 using relative quantification based on the 2nd derivative maximum method. Each cDNA sample was assayed in triplicate. The primers used for *ACTIN7* (*ACT7*) were 5'–CAGTGTCTGGATCGGAGGAT–3' and 5'–TGAACAATCGATGGACCTGA–3'; for *CO* were 5'–TAACAGTAACACAACTCAGTCC–3' and 5'–CCTCGAAGCATACCTTATTGTC–3'; and, for *FT* were 5'–GATCCAGATGTTCCAAGTCC–3' and 5'–ACAATCTCATTGCCAAAGGT–3'. Transcript levels were normalized to *ACT7* expression [192]. Expression analysis of *CO* and *FT* transcription in *35S:3HA-CO* constructs has been previously described [153].

- *Immunoblot analysis and protein quantification*

To detect CDF1 protein in *35S:HA-CDF1* [90] and *35S:HA-CDF1*;*gi-2* [81] and CO protein in *35S:3HA-CO* transgenic lines, plants from the Col-0 (Columbia) accession were grown on Linsmaier and Skoog (LS) media (Caisson, Rexburg, Idaho, USA) containing 3% sucrose at 22°C or 27°C with a fluence rate of 60 μmol m^{-2} s^{-1} in long-day (16L:8D) and short-day (8L:16D) conditions for 10 days. Seedlings were harvested at each time point on day 10 and were ground in liquid nitrogen for protein

extraction. Whole proteins including the nuclear fraction were extracted with buffer containing 50 mM Tris, pH 7.4, 100 mM KCl, 10% glycerol, 5 mM EDTA, 1.0% NP-40, 0.5% deoxycholate, 0.1% SDS, 50 μM MG-132 and Complete protease inhibitor cocktail tablets (Roche, Indianapolis, Indiana, USA). Approximately 50 ug of extracted proteins were resolved in 12% SDS-PAGE gels, and transferred to Nitrocellulose membranes (Whatman, GE Healthcare, USA). HA-CDF1 protein was detected using anti-HA HRP conjugated (3F10, Roche, Indianapolis, Indiana, USA) and visualized with SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific, USA). For quantification of HA-CDF1 protein, non-specific binding of anti-HA around 25 kDa was used as a loading control. The method for protein quantification was described previously [153].

- *Chromatin Immunoprecipitation (ChIP) assay*

Chromatin immunoprecipitation was performed following the protocol in [193] with modifications using plants expressing the pCCA1:CCA1-YFP protein construct that has been described previously [194]. Wild-type seedlings from the Col-0 (Columbia) accession were grown on ½ MS agar plates at 22°C for 14 days with 12L:12D cycles and harvested at ZT2. The chromatin was sheared to between 100 and 1,000 bp in a Bioruptor UCD 200 (Diagenode, Liege, Belgium) at high intensity for 10 minutes (cycles of 30s on / 30 s off) at 4°C after [195]. An aliquot of the chromatin was reserved at this point as the Input chromatin. Immunoprecipation used equilibrated Dynabeads Protein A (Invitrogen/Life Technologies, Paisley, UK). The pre-cleared chromatin was transferred away from the beads and incubated with rotation over night at 4°C with a 1:1000 dilution of anti-GFP (Abcam ab290; Abcam, Cambridge, UK). A new aliquot of equilibrated beads was then added and incubated with the chromatin solution for 2 hours at 4°C with rotation and then washed with low salt, high salt, and lithium chloride washes. The immunocomplexes were recovered from the beads by boiling for 10 minutes in the presence of 10% Chelex resin (BioRad, Hemel Hempstead, UK) and the proteins removed using Proteinase K Solution (Invitrogen/Life Technologies, Paisley, UK) at 50°C. The reserved Input chromatin was also processed in parallel with Chelex and Proteinase K and then purified using

QIAquick PCR purification Kit (Qiagen, Manchester, UK). qPCR on the ChIP and Input DNA was performed in triplicate using Brilliant III Ultra-Fast SYBR Green QPCR Master Mix (Agilent, Wokingham, UK) on a Mx3005P machine. The results were calculated so that percent input was equal to $100*(primer\_efficiency^{dcT})$ where dcT is the difference between the adjusted input cT and the ChIP sample cT. The input cT was adjusted to account for the dilution factor of the input chromatin. The primer efficiency was unique to each primer pair and is equal to $10^{\frac{-1}{slope(cT=m*\log(Input\_concentration+b)}}$. Primers used in this study were *ACT7* (AT5G09810: 5' – GTATCGGGTGACAATGCAGCTATTA – 3' and 3' – TGCTGGAGTAAAACATAAGCCACTC – 5'), *FKF1* (AT1G68050: *FKF1-a*, 5' – CGAGAATCGCGTTTCACAAA – 3'and 3' – AATATCCCCTGGTGACGTGT – 5'; *FKF1-b*, 5' – ACGAAAATTGCCACCAACTC – 3' and 3' – AAAATGGCGAGAGAACATGC – 5'; *FKF1-N*, 5' – GATTGCAGGGCTTCACTCTC – 3' and 3' – CGTCATGGAGGATCCTGAAT – 5') and *CDF1* (AT5G62430: *CDF1-a*, 5' – CGCGATGCTGACATTTACCT – 3' and 3' – ATTGCATCCTCGTAGGAGCA – 5'; *CDF1-b*, 5' – TTGGTTAACGGAAAGTTTAGTGA – 3' and 3' – CACGATTTCCAAACCAAAA – 5'; *CDF1-N*, 5' – TGGACAACACTTGGGATCAA – 3' and 3' – TGTTCTTTGTGCAAACCCTG – 5'). The *GI* primers, used as a positive control, are the same as those in [179].

Data analysis for modelling

Data used in this chapter was either newly obtained for this study or taken from published sources. As with the previous chapter, published data for *CDF1*, *CO* and *FT* mRNA levels across all conditions available were normalised to the WT peak of *FT* mRNA in LD conditions (see Chapter 3; [50, 74, 78]). Due to the lack of experimental data of *FKF1* mRNA in LD conditions, data was normalised to the peak of WT *FKF1* expression from SDs [78]. This means that *FKF1* mRNA levels are not directly comparable with other components in the system. As with the

previous chapter, FKF1 and CDF1 protein levels were normalised to the WT peak of LD FKF1 expression (see Chapter 3; [153]).

Model derivation

As discussed in Chapters 1 & 2, more recent versions of the circadian clock model are able to describe a wider range of photoperiods and genetic perturbations through over-expression or removal of clock components. In order to describe as many conditions as possible, the model presented here has been developed from the 'Pokhilko2011' circadian clock model [12]. This system includes the transcription repressing EC protein complex [12]. However, in this newer model the simulated peak of *TOC1* mRNA, the component that was used as a 'hidden' activator of *CO* transcription in the previous chapter, does not coincide with dusk across photoperiods [12, 80]. Thus, a different component from the circadian clock model, namely $COP1_n$, has been used as this hidden activator (termed NP = Night-time Peak) as modelled activity of $COP1_n$ is constrained to the dark periods of L:D cycles [12, 211]. The model is shown schematically in Figure 4.1.

The model equations for the flowering pathway are:

$$\frac{dc_{CDF1}^{(m)}}{dt} = \left(n_1 + n_2 \frac{c_L^a}{g_1^a + c_L^a} \right) \frac{g_2^b}{g_2^b + (c_{P9} + c_{P7} + c_{NI} + c_T)^b} - m_1 c_{CDF1}^{(m)} \tag{4.1}$$

$$\frac{dc_{CDF1}}{dt} = p_1 c_{CDF1}^{(m)} - m_2 \left(1 + p_2 c_{FKF1} c_{Gn} + p_3 c_{Gn} - p_4 \frac{c_{E3n}}{k_1 + c_{E3n}} \right) c_{CDF1} \tag{4.2}$$

$$\frac{dc_{FKF1}^{(m)}}{dt} = q_1 L c_P + n_3 \frac{g_3^c}{g_3^c + c_L^c} \frac{g_5}{g_4 + c_{EC}} - m_3 c_{FKF1}^{(m)} \tag{4.3}$$

$$\frac{dc_{FKF1}}{dt} = p_5 c_{FKF1}^{(m)} - p_6 \left(m_4 - L \frac{c_{Gn}}{k_2 + c_{Gn}} \right) c_{FKF1} \tag{4.4}$$

$$\frac{dc_{CO}^{(m)}}{dt} = B_{CO} + \frac{g_5^d}{g_5^d + c_{CDF1}^d}\left(n_4 + n_5 D\frac{c_{COP1n}}{g_6 + c_{COP1n}}\right) - m_5 c_{CO}^{(m)} \tag{4.5}$$

$$\frac{dc_{CO}}{dt} = p_7 c_{CO}^{(m)} - p_8\left(m_7 + m_6 D c_{COP1n} - L\frac{c_{FKF1}}{k_3 + c_{FKF1}}\right)c_{CO} \tag{4.6}$$

$$\frac{dc_{FT}^{(m)}}{dt} = \left(n_6 + n_7\frac{c_{PIF}}{g_7 + c_{PIF}}\right)\left(n_8 + n_9\frac{g_8}{g_8 + c_{CDF1}}\right)\frac{c_{CO}^e}{g_9^e + c_{CO}^e} - m_8 c_{FT}^{(m)} \tag{4.7}$$

where the model parameters take on the same notation as in Chapter 3 and the circadian clock model [12]. Hence, n_i and p_i represent transcription rates and translational modulation respectively; m_i are the degradation rates; g_i and k_i are the Michaelis-Menten constants interpreted as binding affinities; a, b and c are the Hill coefficients; B_{CO} is the basal transcription rate of *CO* mRNA; q_i represents acute light activation of transcription, as in the circadian clock model; L and D represent light and dark from the light function present in the model of the circadian clock, connecting the system to the photoperiod. Superscript (m) highlights mRNA components of the model, other components represent protein levels. c_{PIF} represents total PIF protein levels as determined by a model of photoperiodic hypocotyl elongation developed by Dr. Daniel Seaton, a collaborator from Prof. Andrew Millar's group (University of Edinburgh, UK; Appendix B). The model equations for the photoperiodic hypocotyl elongation model take into account the results of Chapter 2 such that *PIF* mRNA levels are regulated by CCA1/LHY and the EC (Figure 4.1). In the results section of this chapter, I shall discuss the reasoning behind the choices of circadian regulation for *CDF1* and *FKF1* mRNA and protein, as well as PIF regulation of *FT* mRNA.

To calculate the days taken for a plant to flower from the expression level of *FT* mRNA, the same function that was used in [82] was used here:

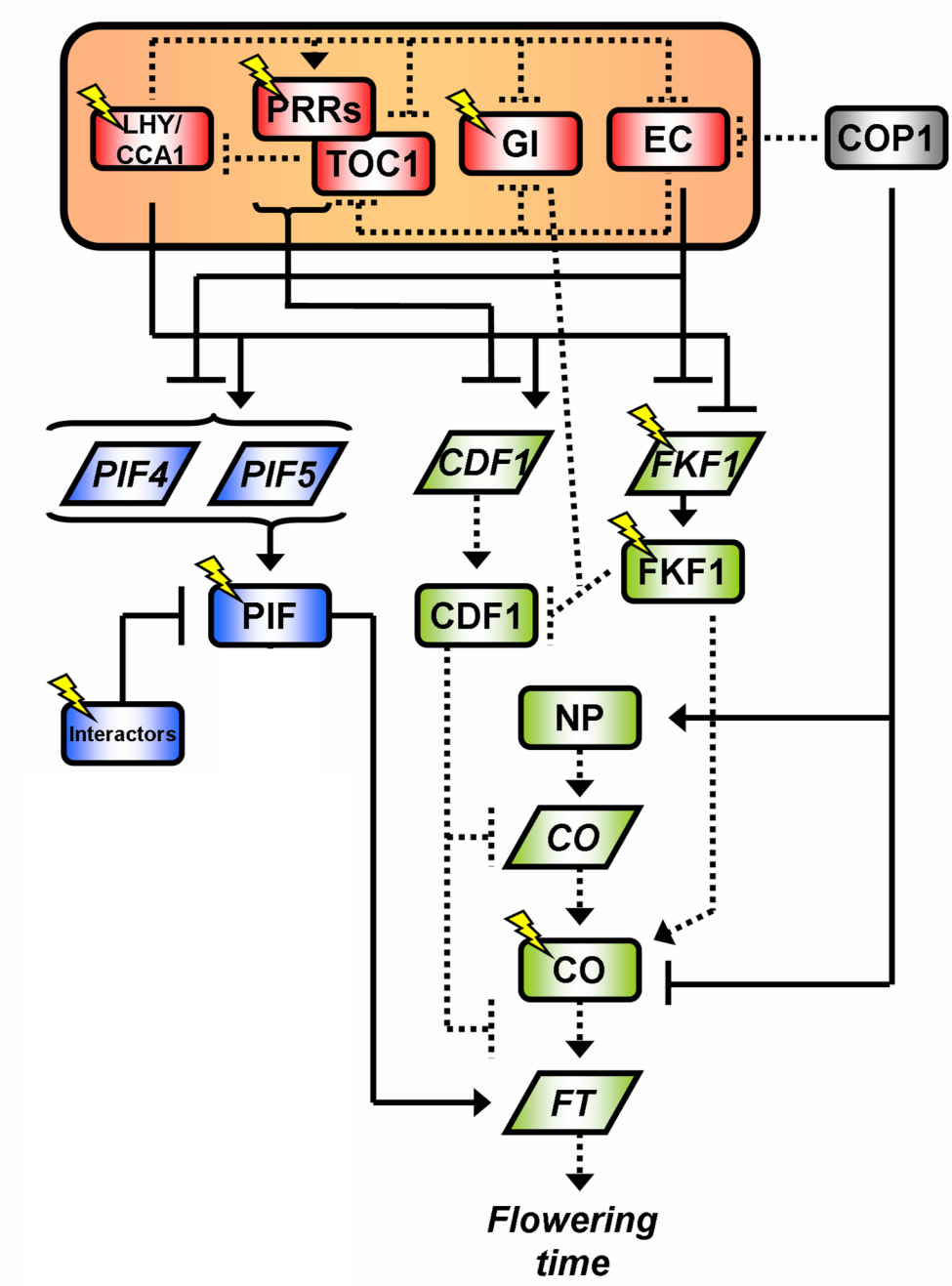


Figure 4.1: Schematic of flowering model. The model of flowering presented in this chapter is shown schematically where the 'Pokhilko2011' circadian clock model is shown in orange and red; the flowering time model presented in Chapter 3 is shown in green with dashed connections; a model of *PIF4/5* regulation is shown in blue and has similar circadian regulation as discussed in Chapter 2 (Appendix B). Parallelograms represent mRNA components, rectangles represent proteins. Thunderbolts represent light-regulated components. NP = night-time peak activator of *CO* mRNA; Interactors = group of proteins that interact directly with PIF proteins (see Appendix B).

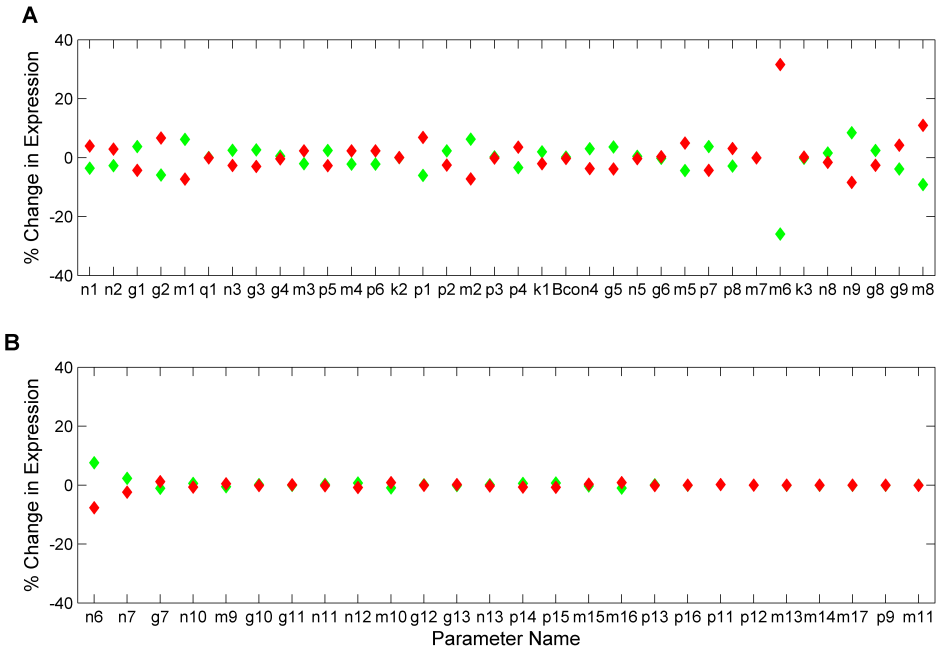


Figure 4.2: Sensitivity of *FT* mRNA simulations to changes in parameter values. The percentage change in FT_{AREA} from LD simulations is recorded for a 10% increase (green diamonds) and decrease (red diamonds) compared to the optimal parameter set used in this study. (A) Percentage change in FT_{AREA} for parameters in photoperiodic flowering pathway (green in Figure 4.1). (B) Percentage change in FT_{AREA} for parameters regulating *FT* mRNA through PIF pathway (blue in Figure 4.1; Appendix B).

$$Days\_to\_Flower = d_0 + \frac{a_F}{1 - FT_{AREA} / b_F} \tag{4.8}$$

where FT_{AREA} is the integrated amount of *FT* expression from a single diurnal cycle (as in Chapter 3). The parameter values for this equation are also provided in Appendix B.

Computational Methods

The model equations for CDF1 protein, *CO* mRNA, CO protein and *FT* mRNA were fitted to data from WT and the *fkf1* loss-of-function mutant, as in the previous chapter [82, 153]. The parameters describing PIF regulation of *FT* mRNA were further fitted to newly acquired data in *pif4;5* loss-of-function mutants under LD conditions (K. Stewart, G. Steel, J. Foreman, University of Edinburgh, unpublished). Equation (4.1), representing *CDF1* mRNA, was optimised to WT and *cca1;lhy* loss-

of-function data in LD conditions [50]. Equation (4.3), representing *FKF1* mRNA, was fitted to data from WT and *cca1;lhy* plants in SD conditions [78]. Additionally, acute light activation of *FKF1* was estimated directly using data from [142]. FKF1 protein dynamics (equation 4.4) included the observations that GI and FKF1 are able to form a blue light-dependent protein complex that stabilises FKF1 protein [81, 83]. Parameters for this equation were fitted to WT data used in Chapter 3 with the observation that the peak expression level of FKF1 protein in the *gi* loss-of-function mutant is approximately half that of WT [83, 153]. In total, this meant that 25 parameters were fitted directly or indirectly to multiple data sets, whilst 5 parameters were constrained by published observations of the system.

As in Chapter 3, the model parameters were fitted to data using the simulated annealing algorithm simulannealbnd from the Matlab R2008b Optimization Toolbox (Mathworks, Cambridge, UK). The parameter values used in this study are given in Appendix B. As with in Chapter 3, we conducted a local sensitivity analysis of the parameter values by measuring the change in *FT* mRNA expression resulting by an increase or decrease of individual parameter values by 10% (Figure 4.2). The high sensitivity of the system to parameter m_6, that describes the dark- and COP1-dependent degradation of CO protein, is due to this rate being the fastest in the system by a factor of 10 or more (Appendix B). As such, perturbing the value of m_6 by 10% leads to the greatest changes in *FT* mRNA levels compared to altering the other, relatively much smaller, parameters. The same result was observed from the model presented in Chapter 3 (Figure 3.7).

Results

Mechanisms for the circadian regulation of CDF1 and FKF1 transcription

As discussed in the introduction to this chapter, a number of datasets have been published from which I can elucidate approximate mechanisms that the circadian clock utilises to regulate *CDF1* and *FKF1* transcription.

- *CDF1 transcription requires more than PRR inhibition*

Most of the experimental evidence from analysis of transcript rhythms in transgenic lines and promoter activity from ChIP-qPCR experiments suggests that *CDF1* transcription is inhibited by the PRR family of proteins [31, 50, 78]. For example, compared to the WT rhythm of *CDF1* levels that peaks ~4hrs after dawn, a *prr9;7* loss-of-function mutant leads to continued expression of *CDF1* mRNA throughout the day (Figure 4.3e; [50]). However, the maintenance of *CDF1* rhythms in *prr9;7* suggests that further circadian clock components are involved in *CDF1* transcriptional regulation. Interestingly, rhythms of *CDF1* mRNA in *cca1;lhy* loss-of-function transgenic plants have ~3hr phase advance compared to WT with a lower mean level of expression (Figure 4.3a; [50]). According to the 'Pokhilko2011' circadian clock model, though, similar levels (i.e. very low) of PRR proteins should be present in simulations of both *prr9;7* and *cca1;lhy* mutants (see clock model schematic in Figure 4.1; [12]) and would, therefore, result in similar phenotypes of *CDF1* mRNA expression if *CDF1* transcription is solely regulated by PRRs. This goes against the data of Figures 4.3a & e where the effects of *prr9;7* and *cca1;lhy* loss-of-function mutants show opposing effects on *CDF1* mRNA expression. This counter-intuitive result suggested that CCA1/LHY may play a role in activating transcription of *CDF1* in conjunction with repression from the PRR family. Modelling such a system, using *CDF1* mRNA data from WT and a *cca1;lhy* loss-of-function mutant (Figure 4.3a), provided qualitatively accurate rhythms of *CDF1* mRNA in simulations of the *prr9;7* double mutant (compare Figure 4.3e with Figure 4.3i).

- *FKF1 mRNA is regulated in a similar manner to GI*

As discussed above, transcriptional control of *FKF1* by the circadian clock seems to correlate closely with regulation of *GI*. The two components share the same phase at around ~9-10hrs after dawn and react in a similar manner to perturbations of the circadian clock [78, 80, 209]. Furthermore, both components are acutely activated by red light [23, 142]. By modelling *FKF1* mRNA in the same way to *GI* transcription

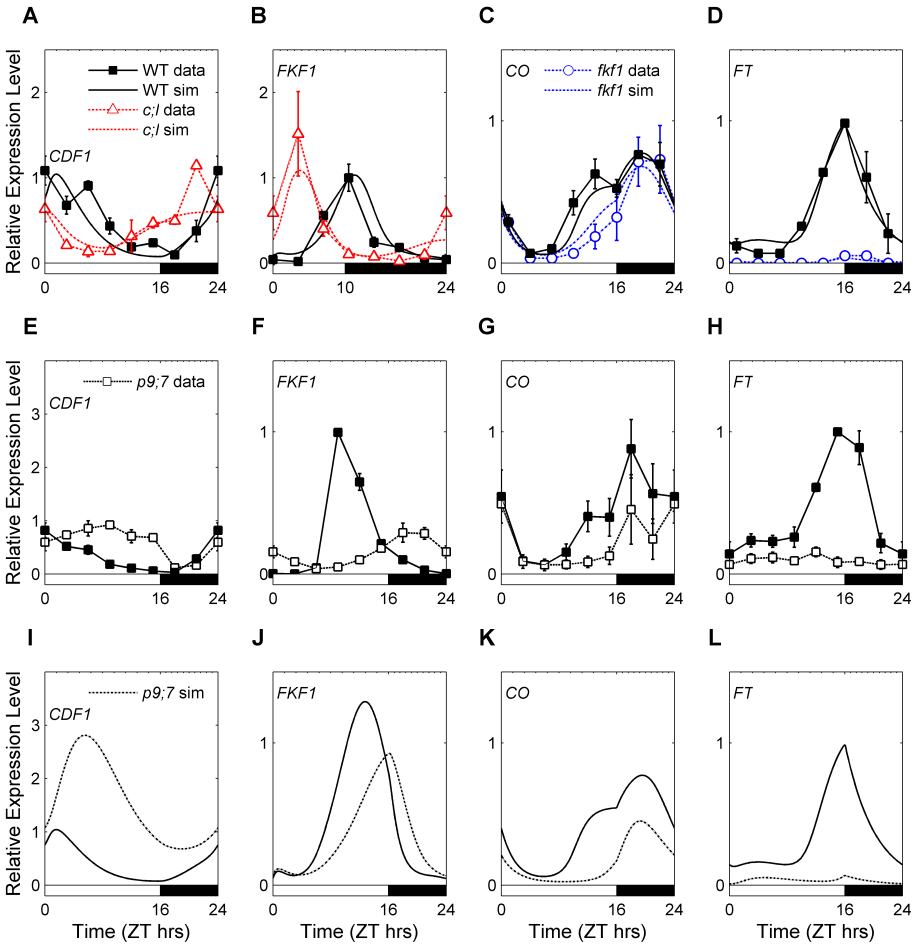


Figure 4.3: Simulations of model components qualitatively match data from a loss-of-function *prr9;7* transgenic line. (A-D) Simulations of (A) *CDF1* mRNA in LD cycles, (B) *FKF1* mRNA in SD, (C) *CO* mRNA and (D) *FT* mRNA in LD compared to training datasets used for optimisation. WT data = solid black lines, filled squares; *cca1;lhy* data = dashed red lines, empty triangles; *fkf1* data = blue dashed lines, empty circles. Simulations for WT = solid black line; *cca1;lhy* = dashed red line; *fkf1* = dashed blue line. Data taken from (A) [50]; (B) [78]; (C, D) [74]. (E-H) As in (A-D) for WT and *prr9;7* data from LD cycles. *prr9;7* data = dashed black lines, empty squares. Data taken from [50]. (I-L) Corresponding simulations of WT and *prr9;7* to compare with (E-H). *prr9;7* simulation = dashed black lines. Data is replotted from published figures, error bars of (A, B, E-H) are standard deviation, (C, D) standard error.

(albeit with different parameter values) such that transcription is repressed by both CCA1/LHY and EC, the model for *FKF1* mRNA was able to provide a close match to the training data (WT and *cca1;lhy* mutant rhythms in SD; Figure 4.3b) and qualitatively similar rhythms to *FKF1* transcription in an LD grown *prr9;7* loss-of-function transgenic plant (compare Figures 4.3f & j). Such a match to the data

confirmed that this mechanism includes key features of the 'true' biological system the model was constructed to describe. As both *CDF1* and *FKF1* transcription can be simulated with this model, the use of data sets of *CDF1* mRNA and FKF1 protein is negated. Previous flowering models, such as the system presented in Chapter 3, required these data sets as inputs to the model so that downstream components could be simulated. Thus, this model now enables users to simulate the flowering system in multiple photoperiods and circadian clock-related transgenic plants without the need of acquiring data beforehand.

- *ChIP assay shows CCA1 activity on the promoters of CDF1 and FKF1*

To confirm the role of CCA1 in the regulation of *CDF1* and *FKF1* transcription, a chromatin immunoprecipitation (ChIP) assay was conducted to determine whether CCA1 is active on their promoters (as in Chapter 2, this is defined as either the 3000bp before the transcription start site or to the nearest gene; Figure 4.4a; D.R. MacGregor, University of Exeter, unpublished). The previously published observation of CCA1 activity on the *GI* promoter was confirmed as in Chapter 2 [179]. By comparing CCA1 activity at CBS sites in the promoters to exons in the respective genomic regions of *CDF1* and *FKF1*, the results confirm that CCA1 is found to be significantly active on the promoters of *CDF1* and *FKF1* to regulate transcription (Figure 4.4b). Thus, the hypothesis that CCA1 is important for the correct regulation of *CDF1* and *FKF1* obtained from the modelling above is confirmed experimentally.

Regulation of CO and FT transcription by the circadian clock

Now that the model accurately described the circadian regulation of *CDF1* and *FKF1* mRNA and could be simulated in multiple conditions, I wished to determine whether simulated rhythms of *CO* and *FT* mRNA provided a qualitatively good match with data from circadian clock mutants. The earlier model could not be tested in this way due to the lack of appropriate input data sets, particularly FKF1 protein, in these

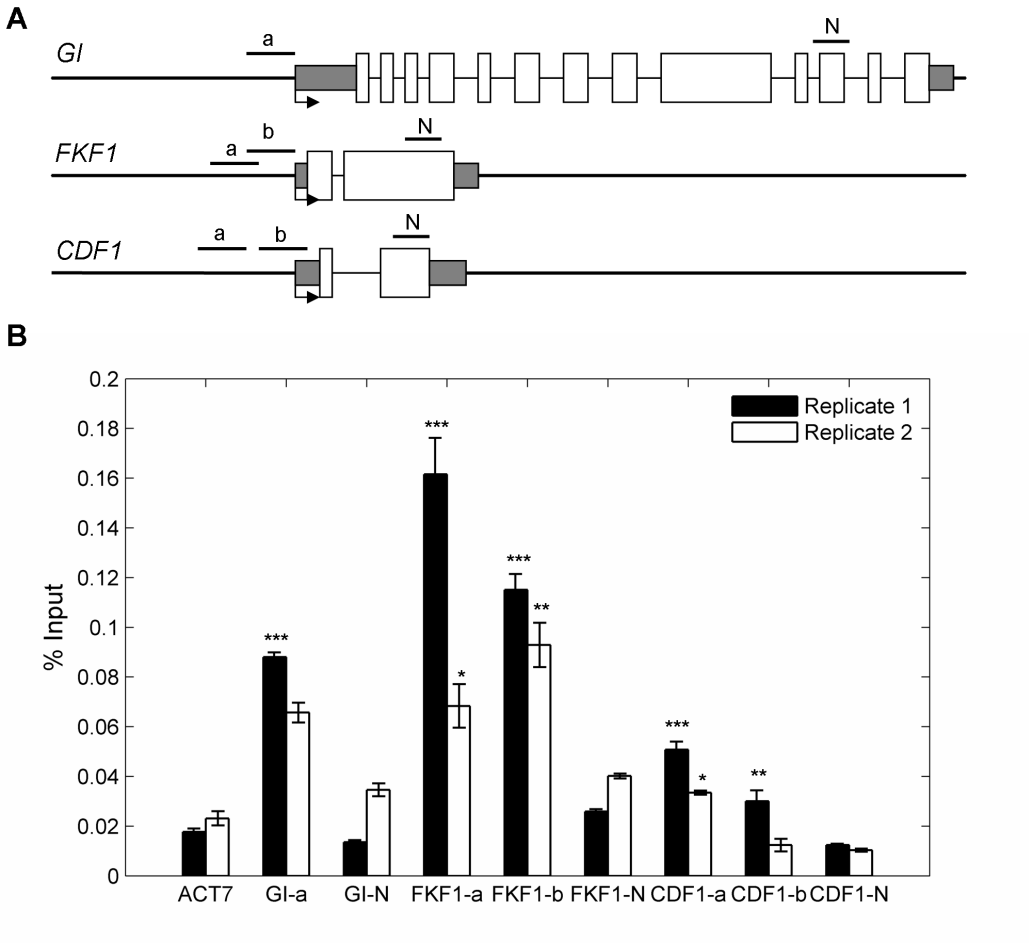


Figure 4.4: CCA1 is active on the promoters of *CDF1* and *FKF1*. (A) Schematic of genomic regions of *CDF1*, *FKF1* and *GI* that is used as a positive control. All regions in the promoter (*a* or *b*) confer the presence of CBS sites. Sites *N* are negative controls in the exons of each gene. (B) ChIP assay shows significant CCA1 activity at the CBS binding sites compared to the respective exons in two biological replicates. Welch tests were used for statistical analysis: (*) $p < 0.1$; (**) $p < 0.05$; (***) $p < 0.005$. Error bars represent standard error.

backgrounds. The motivation was to determine whether the circadian regulation of *CDF1* and *FKF1* transcription was sufficient to explain *CO* and *FT* mRNA profiles or whether further circadian-regulated components are missing from the system presented here. The model equations that describe *CO* and *FT* transcription were fitted to WT and an *fkf1* loss-of-function mutant (Figure 4.3c & d). As the models of *CDF1* and *FKF1* mRNA were able to describe *cca1;lhy* and *prr9;7* loss-of-function mutants in LDs, I compared simulation results of *CO* and *FT* mRNA in these backgrounds to published data. As seen in Figure 4.3k & l, simulations of *CO* and *FT* mRNA in the *prr9;7* double mutant qualitatively match the data presented in Figure

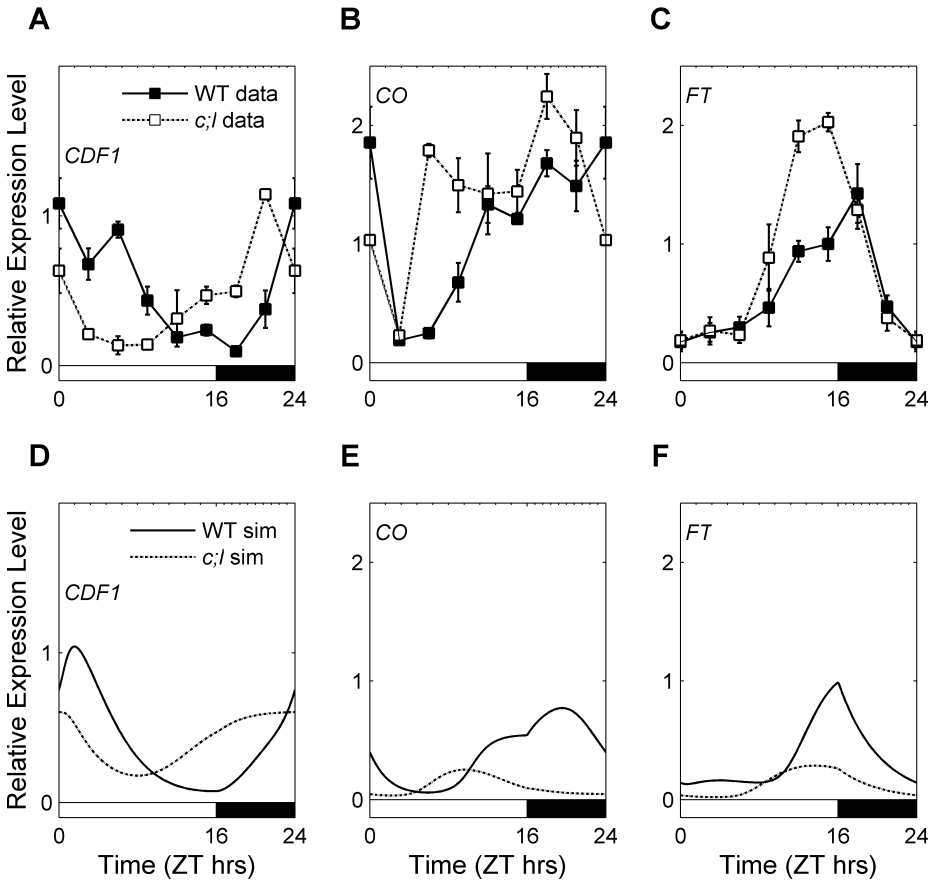


Figure 4.5: Flowering model is unable to describe *CO* and *FT* mRNA in the *cca1;lhy* loss-of-function double mutant. Model simulations qualitatively compared to data for components of the flowering pathway under LD conditions. Data from WT are represented with solid lines and filled squares; *cca1;lhy* mutant are dashed lines and empty squares. Simulations are similarly represented with solid line for WT and dashed line for *cca1;lhy*. (A and D) Data and simulation of *CDF1* mRNA. (B and E) Similar to (A, D) for *CO* mRNA. (C and F) Same for *FT* mRNA. Data taken from [50]. Error bars represent standard deviation.

4.3g & h [50]. Thus, it appears that regulation of *CDF1* and *FKF1* transcription is sufficient to describe the effects of PRR9 and PRR7 on the flowering pathway. Interestingly, though, LD simulations of *CO* and *FT* mRNA in *cca1;lhy* loss-of-function mutants were unable to provide a similar qualitative match with published datasets (compare Figures 4.5b & c with 4.5e & f). Simulations of both *CDF1* and *FKF1* mRNA are phase advanced in the *cca1;lhy* mutant background (Figure 4.3a & b and Figure 4.5a), resulting in an advanced phase of *CO* transcription in LDs (Figure 4.5e). However, our model is unable to match the continued increase of *CO* transcription during the second half of the day and the night period of LD diurnal

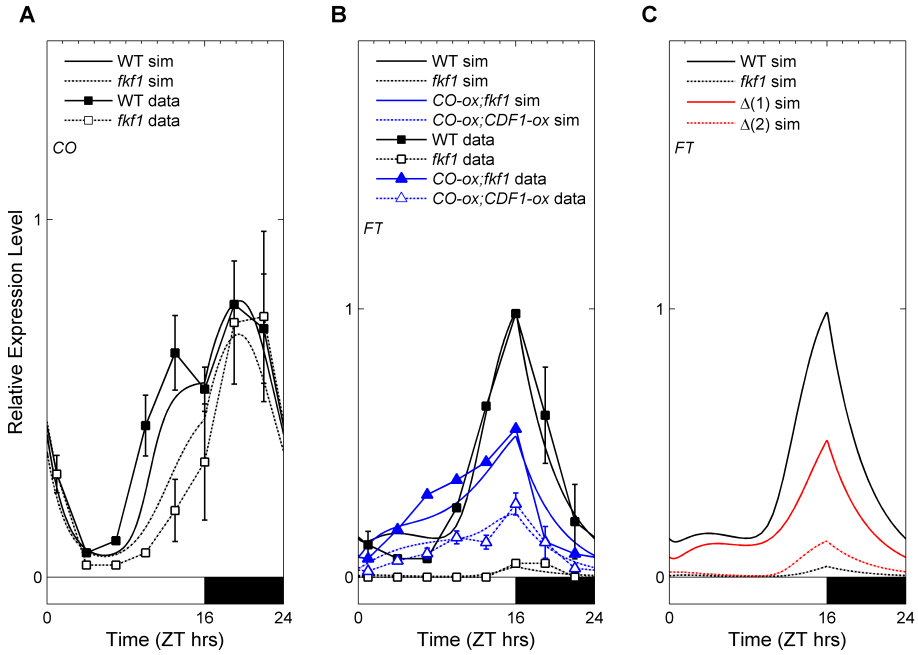


Figure 4.6: Comparison of results from previous flowering time model with new model. Key results of previous flowering model are recapitulated with LD simulations from the updated model presented here (Chapter 3, [153]). (A) *CO* mRNA simulated in WT (solid line) and *fkf1* mutant (dashed line) are compared to data (WT = solid line, filled squares; *fkf1* = dashed line; empty squares). (B) *FT* mRNA in WT, *fkf1*, *CO-ox;fkf1* and *CO-ox;CDF1-ox* (error bars for this line were unavailable) lines are simulated and compared to data. WT and *fkf1* lines are represented as in (A). *CO-ox;fkf1* simulations = blue solid lines; data = blue solid lines; blue-filled triangles. *CO-ox;CDF1-ox* simulations = blue dashed lines; data = blue dashed lines; empty triangles. (C) *FT* mRNA in WT, *fkf1*, Δ(1) and Δ(2) *fkf1*-partial mutants (Chapter 3, [153]). Δ(1) = red solid line; Δ(2) = red dashed line. WT and *fkf1* as in (A). Data from WT and *fkf1* lines taken from [74]. Data from *CO-ox;fkf1* and *CO-ox;CDF1-ox* taken from [153]. Error bars represent standard error.

cycles (compare Figure 4.5b & e). This discrepancy between the data and simulations of *CO* mRNA leads to a similar discrepancy of *FT* mRNA, such that simulated *FT* transcription starts to increase earlier in the day but does not continue into the second half of the day as seen in the data (compare Figure 4.5c & f). Since the simulations of *CDF1* mRNA (Figure 4.5d) qualitatively match the data presented in Figure 4.5a, one could hypothesise that CCA1 & LHY plays a further role in the flowering system between *CDF1* and *CO* transcription that is currently not described by the model presented here. These hypotheses will be elaborated on in the discussion section of this chapter.

Model predicts a novel CDF1-ELF3 interaction to describe the key flowering loss-of-function gi mutant

To ensure that the model described in this chapter was consistent with the model presented in Chapter 3, a number of the key flowering mutations were simulated. As shown in Figure 4.6a & b, the transcriptional rhythms of *CO* and *FT* were able to be described in the *fkf1*, *CO-ox*, *CO-ox;fkf1* and *CO-ox;CDF1-ox* simulations as in the model of Chapter 3. Furthermore, in Figure 4.6c, the model constructed here was able to match the conclusions drawn in Chapter 3, namely that the CO-FKF1 interaction has a relatively larger role than the CDF1-FKF1 interaction in determining the correct levels of *FT* expression (this is shown by $\Delta(1)>\Delta(2)$, see Chapter 3 for details). Interestingly, though, the key flowering *gi* loss-of-function mutant was not well described by the model of *CO* and *FT* mRNA (compare the dashed red line of Figure 4.7c & d with g & h, respectively). The high amplitude rhythms of simulated *CO* mRNA correlate with the low amplitude rhythms of CDF1 protein (Figure 4.7f) despite there not being a great change in *CDF1* transcription between WT and *gi* simulations (Figure 4.7e). Thus, one of two possible mechanisms could occur that would result with minimal simulated transcription of *CO* in *gi* mutants, as described by the data (Figure 4.7c & g): either CDF1 inhibition of *CO* mRNA is increased in the absence of GI such that low levels of CDF1 have increased activity, or; CDF1 protein maintains the same strength of inhibition of *CO* transcription in WT and *gi* mutants, but CDF1 protein levels are increased. In the model of *CO* transcription presented in Chapter 3, the *gi* loss-of-function mutant was accurately described as GI played a role in directly regulating *CO* mRNA in conjunction with CDF1 (see equation 3.2). However, due to the lack of quantified data it was unknown whether the model correctly described the regulation of CDF1 protein in *gi* loss-of-function transgenic plants. Therefore, to determine whether CDF1 protein levels of activity were increased in the absence of GI function an experiment conducted by Dr. Young Hun Song (University of Washington, USA, unpublished) obtained quantified protein levels of CDF1 in a *35S:HA-CDF1;gi-2* transgenic plant (labelled *CDF1-ox;gi* in Figure 4.7b; [81]). The data of Figure 4.7b

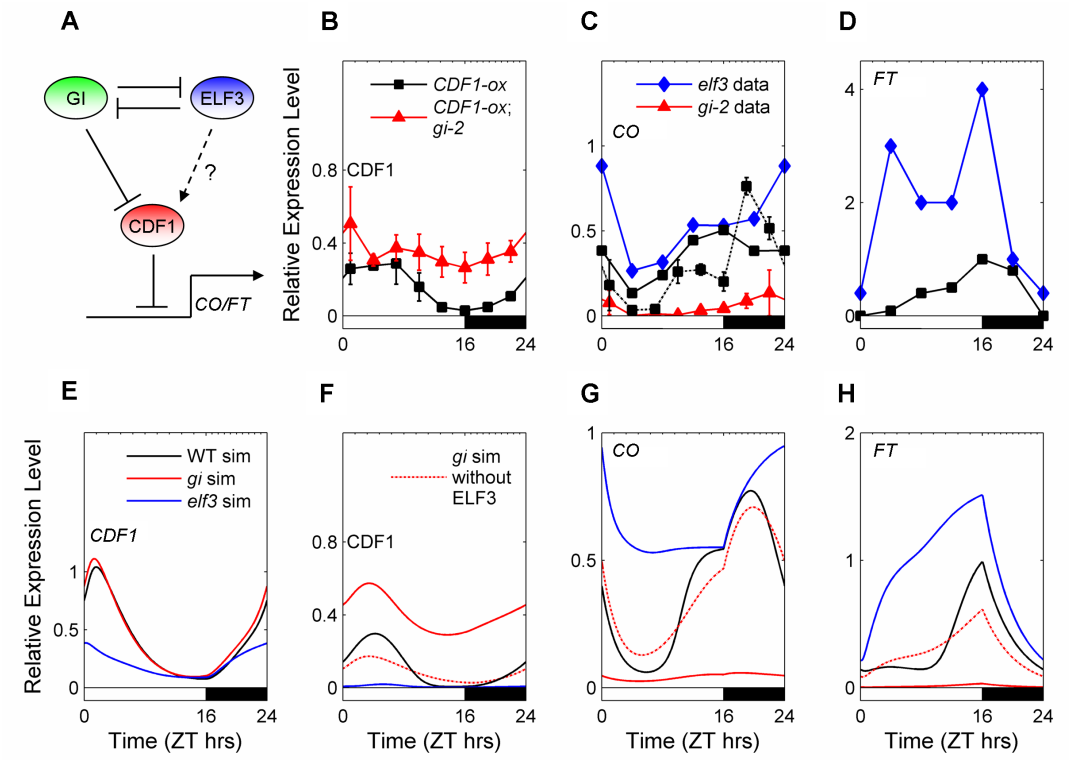


Figure 4.7: ELF3 regulation of CDF1 protein required for correct qualitative description of *gi* mutant. (A) Schematic of ELF3-GI-CDF1 interaction. GI and ELF3 antagonise each others stability and function. GI degrades CDF1 protein, whilst the model proposes that ELF3 stabilises CDF1 levels. CDF1 then inhibits transcription of *CO* and *FT*. (B) CDF1 protein data in 35S:HA-CDF1 (*CDF1-ox*; black lines, filled squares) and 35S:HA-CDF1/*gi-2* mutant (*CDF1-ox/gi-2*; red lines, filled triangles) in LD. (C) *CO* mRNA data from LD in WT (black line, filled squares), *gi-2* (red line, filled triangles) and *elf3* (blue line, filled diamonds). Data from [77] = solid line (error bars were unavailable); [81] = dashed line. (D) *FT* mRNA in WT and *elf3* as in (C). (E) Simulations of *CDF1* mRNA from LD in WT (black line), *gi* (red line) and *elf3* (blue line) mutant. (F-H) Simulations of (F) CDF1 protein, (G) *CO* and (H) *FT* mRNA from LD in WT, *gi* and *elf3* as in (E). Red dashed line is simulation of *gi* mutant without the ELF3-CDF1 interaction. Error bars represent standard error.

show that CDF1 protein levels follow a shallow rhythm with a higher mean expression level in the *35S:HA-CDF1;gi-2* transgenic line compared to *35S:HA-CDF1*. This suggests that the low expression level and shallow rhythms of *CO* mRNA in *gi* loss-of-function mutants occur due to increased CDF1 protein levels, rather than activity (Figure 4.7c and equation 4.5). As the model could not describe the increased levels of CDF1 protein in *gi* mutants, this implied that the system was

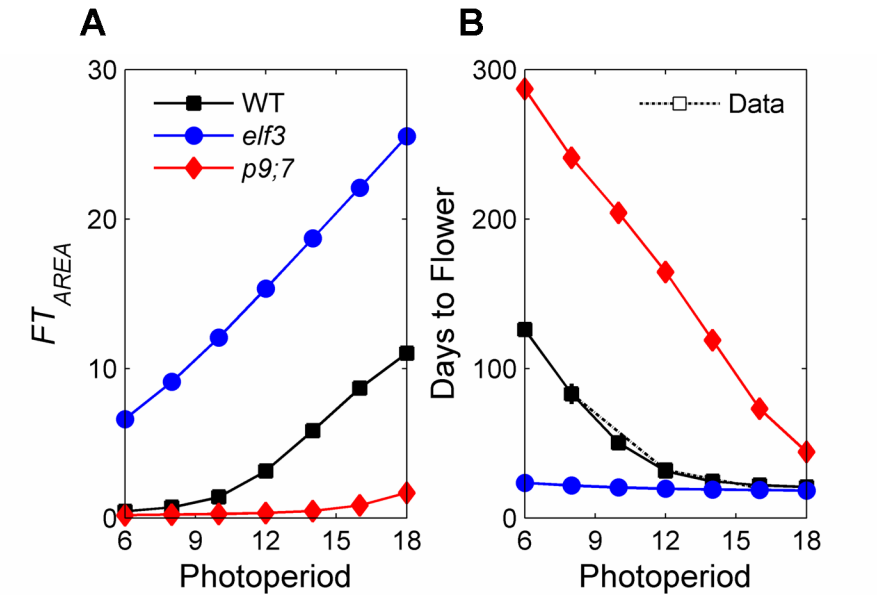


Figure 4.8: Model describes photoperiodic flowering through accumulation of *FT* transcript. (A) FT_{AREA} calculated from simulations of WT (black line, filled squares), *elf3* (blue line, filled circles) and *prr9;7* loss-of-function mutants (red line, filled diamonds) across a range of photoperiods. (B) Using FT_{AREA} from WT simulations, the model is able to match flowering data from [72]. Error bars represent standard deviation.

missing a positive regulator of CDF1 protein, whose role becomes more relevant in the absence of GI protein activity.

A candidate for such a role in this system is ELF3. Studies have shown that ELF3 acts antagonistically against GI, and that the ELF3-GI interaction may be mediated in part by COP1 [12, 34]. Furthermore, manual searching of the *CO* and *FT* promoter amplicons that have been shown to be targeted by CDF1 and GI also contained partial LBS that are currently considered to be required for EC, and ELF3, function on target promoters [25, 28]. These amplicons contained 5/6 nucleotides required for a full consensus LBS sequence (LBS underlined: *CO* – amplicon 9: 5'-GCTA*GATTCG*TTTTATCTCTTTGGC-3'; [81]; and *FT* – amplicon 12: 5'-*GTATC*ATAGGCATGAACCCTCT-3'; amplicon 13: 5'-AGAGGGTTCATGCCTAT*GATAC*-3'; [153]). By adding a CDF1-ELF3 interaction into the dynamics of CDF1 protein (equation 4.2) such that ELF3 stabilises CDF1 protein (as shown schematically in Figure 4.7a), the model simulations were able to qualitatively describe the shallow rhythm, high mean expression levels of CDF1

protein in the *gi* mutant (compare Figure 4.7b & f). As a consequence, *CO* and *FT* mRNA rhythms were correctly described in both the *gi* and *elf3* loss-of-function mutants (Figure 4.7g & h). The results of our simulations suggest that the ELF3-GI interaction may play an important role in flowering through regulation of CDF1 protein levels (Figure 4.7a).

Circadian clock measurement of day-length regulates photoperiodic flowering

The model presented in this chapter is the first full model of flowering that is able to describe the key perturbations to the circadian clock, such as changes in photoperiod and loss-of-function mutations in circadian clock components. Hence, the model is able to accurately match data that shows the days taken for Arabidopsis plants to flower in a number of different photoperiods (Figure 4.8b; [72]). Similarly, simulations of flowering phenotypes from a number of perturbations to the system also match published phenotypes (as seen in Table 4.I). Furthermore, perturbing the circadian clock through *prr9;7* and *elf3* loss-of-function mutants allows the model to predict the effects that removal of these components would have on photoperiodic flowering (Figure 4.8b). Thus, the model shows that removal of ELF3 leads to FT_{AREA} increasing linearly with photoperiod rather than the non-linear manner seen in WT simulations.

Warm, ambient temperatures increase expression of FT mRNA in LDs

As discussed in the introduction to this chapter, under SD conditions, the acceleration of flowering caused by an increase in growth temperature (from ~22°C to ~27°C) is primarily due to PIF4 activation of *FT* transcription in a manner considered to be independent of the CO- and photoperiod-dependent pathway modelled in this chapter. To determine whether the same mechanism regulates temperature dependent flowering in LD, collaborators from Dr. Karen Halliday's group (University of Edinburgh, UK) tested rhythms of *CO* and *FT* mRNA in LD at 22°C and 27°C in both WT and *pif4;5* loss-of-function double mutant plants (Figure 4.9a, b; K. Stewart, G. Steel, J. Foreman, University of Edinburgh, unpublished). As can be seen in Figure 4.9a, rhythms of *CO* mRNA are similar in all the tested

Table 4.I: Predicted flowering phenotypes of transgenic plants in SD (8L:16D) and LD (16L:8D) simulations compared to WT. Phenotypes given in brackets are the phenotypes observed in reference publications. E = Early flowering; D = Delayed flowering; ND = Not Determined.

System Perturbation	Short Days	Long Days	Reference Publication(s)
cca1;lhy	E (E)	D (~WT)	[18]
prr9;7	D (D)	D (D)	[50]
elf3	E (E)	E (E)	[34, 77]
ztl	E (E)	E (~WT)	[77, 110, 212]
gi	D (D)	D (D)	[34, 83]
fkf1	D (D)	D (D)	[90]
CO-ox	E (E)	E (E)	[153]
CO-ox;CDF1-ox	E (E)	D (~WT)	[153]
cdf	E (E)	E (E)	[83]
pif4;5	D (D)	D (ND)	[174]
cop1	E (E)	D (~WT)	[34]

conditions suggesting that PIF4/5 and temperature do not greatly alter the regulation of *CO* transcription. Conversely, in Figure 4.9b, levels of *FT* mRNA from WT plants between ZT8 and ZT16 are greater in plants grown at 27°C compared to 22°C. This effect is partially dependent on PIF4/5, as the rise of *FT* mRNA is less substantial at ZT12 in the *pif4;5* double mutant at 27°C compared to WT levels (Figure 4.9b). This suggests that the temperature response of *FT* mRNA is not completely dependent on PIF4/5, as is evident by the remaining temperature sensitivity of the *pif4;5* double mutant, and indeed the greater temperature induction of *FT* mRNA observed during the early day of *pif4;5* transgenic plants grown at 27°C compared to those grown at

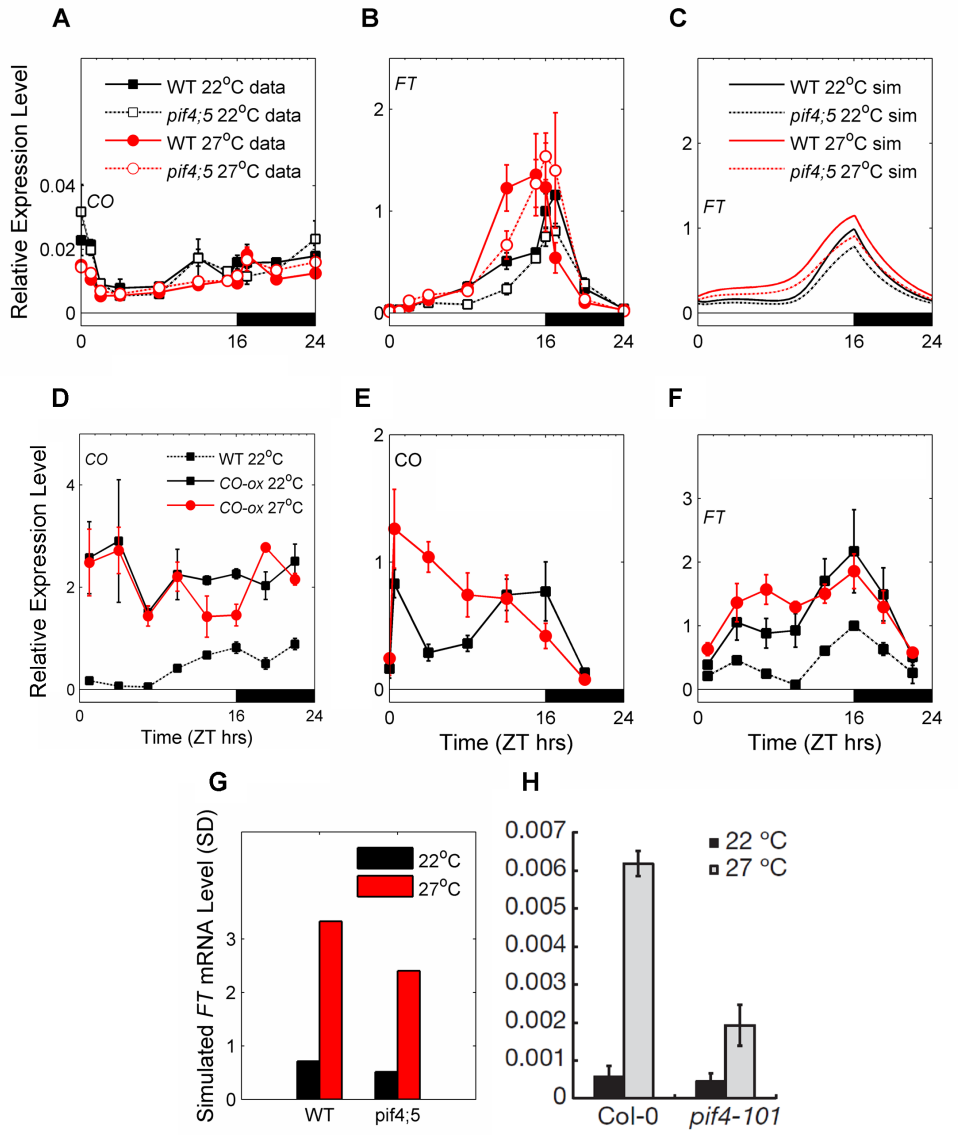


Figure 4.9: PIF4/5 regulates *FT* mRNA with CO protein in LD across temperatures. Data for (A) *CO* mRNA and (B) *FT* mRNA were obtained in WT (22°C: solid black line, filled squares; 27°C: solid red line, filled circles) and *pif4;5* mutants (22°C: dashed black line, empty squares; 27°C: dashed red line, empty circles) under LD conditions. (C) Simulations of *FT* mRNA in WT (solid line) and *pif4;5* mutant (dashed line) in LD at 22°C (black lines) and 27°C (red lines). Simulations of 27°C modelled by $g_9=g_9/10$. (D-F) Data of (D) *CO* mRNA, (E) CO protein and (F) *FT* mRNA in WT (black dashed lines, filled squares) at 22°C and *35S:3HA-CO* at 22°C (black solid line, filled squares) and 27°C (red solid lines, filled circles). (G) Simulated amounts of *FT* transcript produced in SD (8L:16D) conditions in WT and *pif4;5* at 22°C (black bars) and 27°C (red bars), comparable to (H) data from [174] (black bars = 22°C, grey bars = 27°C) in SDs. Error bars represent (A, B, D-F) standard error and (H) standard deviation.

22°C (Figure 4.9b). Moreover, the difference in *FT* mRNA rhythms from WT and *pif4;5* backgrounds between ~ZT4-16 at 22°C shows that PIF proteins also play a limited role in the correct regulation of *FT* mRNA at cooler temperatures in LD conditions (Figure 4.9b).

The data suggested that the temperature regulation of *FT* mRNA occurs in the absence of PIFs specifically between ZT8 and ZT16, the time when the CO-dependent photoperiod pathway is active. Interestingly, the specific timing for temperature regulation of *FT*, and not *CO*, mRNA suggests that temperature regulates flowering downstream of *CO* transcription. A model of *FT* transcription that took into account the role of PIF4/5 in *FT* regulation at 22°C and made levels of CO protein temperature sensitive (decreasing parameter g_9 10-fold between 22°C and 27°C in equation 4.7) allowed the model simulations to qualitatively match the data presented here (compare Figure 4.9b & c). The model derived prediction that CO protein is temperature sensitive was validated experimentally by Assistant Prof. Takato Imaizumi's group (University of Washington, unpublished) who showed that CO protein, in a *35S:3HA-CO* overexpressing transgenic plant, is more stable between ~ZT4-12 of 27°C LDs compared to 22°C LDs despite little change in the mRNA level (Figure 4.9d & e). The higher level of CO protein then corresponds to higher expression of *FT* mRNA at these times in 27°C grown plants compared to those grown at 22°C (Figure 4.9f). Interestingly, simulating the same conditions in SD provided a qualitatively similar result to what has been observed experimentally (Figure 4.9g & h; [174]). In SD, expression levels of *FT* mRNA are increased at 27°C compared to 22°C and this increase is dependent on PIF4 activity, which does not greatly regulate *FT* expression at 22°C. This qualitative similarity to the data supports the view that part of the PIF4 regulation of *FT* transcription does occur through an interaction with the CO-FT pathway.

Discussion

In this chapter I have developed on the model presented in Chapter 3 to achieve two goals: first, to connect *CDF1* and *FKF1* transcription to the 'Pokhilko2011' circadian

clock model and; second, to test the mechanism regulating temperature-induced acceleration of flowering through increased expression of *FT* mRNA. By building a model describing the circadian regulation of *CDF1* and *FKF1* transcription, the model presented here does not require data sets to be input into the system to function, unlike the model of Chapter 3 (see Figure 4.1). Furthermore, this allows for the model to simulate multiple perturbations to the circadian clock system, such as different photoperiods and loss-of-function mutations that have not been able to be simulated by previous models of photoperiodic flowering (Figures 4.3 and 4.5-4.8). The inclusion of model components that are temperature sensitive is also the first mathematical analysis to link temperature with the photoperiodic CO-FT molecular pathway that regulates flowering.

The additional connections of the circadian clock to *CDF1* and *FKF1* transcription were constructed based on the currently available published data sets and observations. Thus, *CDF1* transcription was postulated to be inhibited by the PRR protein family, but also activated by CCA1 & LHY. This mechanism correctly described the rhythms of *CDF1* mRNA from SD and LD conditions in WT, *cca1;lhy* and *prr9;7* plants (Figure 4.3a & i; [31, 50, 78]). Based on the observations that *FKF1* transcription is seemingly under the control of a similar mechanism that regulates *GI* transcription, the model of *FKF1* regulation had transcription inhibited by both CCA1/LHY and EC. This model was able to correctly describe the effects of *cca1;lhy* and *prr9;7* loss-of-function mutants compared to WT rhythms, that were similarly timed to rhythms of *GI* (Figure 4.3b & j; [50, 78, 80]). Furthermore, the regulation of both *CDF1* and *FKF1* transcription by CCA1 was validated experimentally, by showing CCA1 activity at CBS sites lying upstream of their respective transcription start sites (Figure 4.4; D.R. MacGregor, University of Exeter, unpublished). Through these connections the simulations of *CO* and *FT* transcription in the *prr9;7* loss-of-function mutant were qualitatively similar to published data (Figure 4.3k & l; [78]).

Interestingly, though, simulations of *CO* and *FT* mRNA in the *cca1;lhy* loss-of-function double mutant were unable to qualitatively match published data sets

(Figure 4.5e & f). Despite *CDF1* mRNA being described by the presented model in LD conditions in both WT and *cca1;lhy* mutant, rhythms of *CO* mRNA were unable to maintain a high level of expression (Figure 4.5e). The lack of simulated *CO* mRNA in the *cca1;lhy* mutant then leads to an under-estimate of *FT* mRNA compared to the published data that shows a large increase in *FT* expression (compare Figure 4.5c & f). Thus, it seems likely that CCA1 or LHY play a further role in the flowering system by inhibiting *CO* and *FT* mRNA directly. As the model describes *CDF1* mRNA correctly in the *cca1;lhy* mutant, the role of CCA1 or LHY on *CO* and *FT* mRNA must, logically, be either downstream of *CDF1* mRNA or in an independent parallel pathway. In the model presented here this opens up two possibilities. First, CCA1 or LHY might be part of the circadian regulation of CDF1 protein. Interestingly, ELF3 and CCA1 have been observed to interact *in vivo*, and it may be through this interaction that CCA1 regulates CDF1 [212]. Second, CCA1 or LHY might be part of the unknown NP mechanism (see Figure 4.1) that regulates the night-time peak of *CO* mRNA expression. Recently, two protein groups, LOV1 and FBH1/2, have been found to regulate the night-time peak of *CO* mRNA and it may be through these that CCA1 & LHY plays a secondary role in the flowering system [84, 85].

Further to the incorrect rhythms of *CO* and *FT* mRNA in simulated *cca1;lhy* mutants, the model was also unable to accurately describe *CO* and *FT* transcription in *gi* loss-of-function mutants due to low levels of simulated CDF1 protein (Figure 4.7; [73, 81, 83]). Thus, I hypothesised that CDF1 protein levels or activity must be enhanced in *gi* loss-of-function transgenic plants such that *CO* transcription is suppressed. Experimental evidence showed that CDF1 protein levels in the *35S:HA-CDF1;gi-2* line are greater than those in the *35S:HA-CDF1* background (Figure 4.7b; Y.H. Song, University of Washington, unpublished; [81]). Thus, low levels of *CO* mRNA in *gi* loss-of-function transgenic plants are a result of increased levels of CDF1 protein. However, this suggested that the model was missing a component that stabilised CDF1 protein in the *gi* loss-of-function mutant. One candidate for such a role is ELF3 due to the antagonistic relationship between ELF3 and GI [34]. Including this mechanism into the model produced simulations of CDF1 protein and

CO & *FT* transcription in the *gi* loss-of-function mutant that are comparable with the data (Figure 4.7). Thus, the model presented here makes the hypothesis that ELF3 plays a role in the regulation of GI-CDF1 interactions, resulting in high levels of CDF1 protein in the *gi* mutant (Figure 4.7a). To confirm the regulation of CDF1 protein stability by ELF3 further experiments would need to be conducted to test whether a direct interaction occurs between ELF3 and CDF1, whether this interaction is dependent on the absence of GI, and how the diurnal rhythm of CDF1 protein is altered in an *elf3* loss-of-function transgenic plant. Since ELF3 has been shown to interact with phyB, the key red light photoreceptor (see Chapter 1), and play a role in the circadian clock through the EC, this mechanism may present another mechanism through which the circadian clock and light signalling can finely tune the flowering process [12, 25, 214, 215].

As the model presented here is able to describe *FT* transcription in a range of photoperiods and a number of transgenic lines where components of both the circadian clock and flowering pathway are perturbed (Figure 4.8 and Table 4.I), I wished to add the regulation of *FT* mRNA across temperatures. By including the temperature regulation of flowering, this model represents a first step towards constructing an artificial molecular system that may describe flowering in more natural conditions where light and temperature sources are not constant throughout a day. The temperature-induced acceleration of flowering is due to an increase of *FT* mRNA that is partially dependent on PIF4 activity and is thought to be independent of CO function [173, 174]. In order to determine whether this is the case, experiments were conducted in LDs when CO is the key regulator of *FT* transcription (K. Stewart, G. Steel, J. Foreman, University of Edinburgh, unpublished). The results highlighted that *FT* mRNA rhythms are temperature sensitive in the *pif4;5* double mutant, suggesting that another regulator of *FT* transcription is able to mediate the temperature-induced increase in *FT* mRNA levels (Figure 4.9b). In the model presented here, I have hypothesised that CO is a temperature-sensitive regulator of *FT* mRNA such that when plants are grown at 27°C the regulation of *FT* transcription by CO protein is increased. This prediction was validated experimentally by Assistant Prof. Takato Imaizumi's group (University of

Washington, unpublished; Figure 4.9d-f). Further to this, the model also predicted that a CO-PIF interaction would be able to explain a proportion of the temperature response of *FT* mRNA seen in both SD and LD growth conditions (Figure 4.9c, g & h; [174]). To test this hypothesis, experiments would need to be conducted to test whether CO and PIF4 directly interact. Interestingly, PIFs have already been shown to interact with the CCT domain of TOC1 protein, which is also present in the CO protein structure [216, 217]. Thus, it seems reasonable to suggest that PIFs may also regulate CO through the CCT domain. Furthermore, an interaction between CO and PIFs may form part of the light regulation of CO protein that has been shown to be degraded quickly in red light and is under the control of phytochrome signalling [86, 153].

In this chapter, I have introduced a more complete model of the photoperiodic flowering pathway. From this system I have derived hypotheses that ELF3, CCA1 and LHY may play a role in the regulation of CDF1 protein and that CO protein is important for the temperature sensitivity of LD flowering through *FT* transcription. These hypotheses should be tested experimentally in future to enhance our understanding of the circadian- and temperature-regulation of flowering. In the next chapter I shall discuss on-going work as part of the ROBuST project that is looking at specific mechanisms through which red light and temperature directly regulates the circadian clock.

Chapter 5: Red light and temperature signals regulate the circadian clock of *Arabidopsis thaliana*

In the previous chapters I have concentrated on how light and temperature are able to regulate plant growth and development through specific mechanisms that are outputs of the circadian clock. However, as mentioned in Chapter 1, temperature and light signalling also play an important role in regulating the circadian clock itself. Yet, across a range of light intensities and temperatures the period length of the circadian clock is maintained at ~24hrs [8–10]. This process is referred to as 'compensation'. In this chapter, data sets produced by Dr. Karen Halliday's group (University of Edinburgh, UK) will be examined to investigate this problem. Using the most recent mathematical model of the circadian clock, the 'Pokhilko2012' model, I will go on to propose a mechanism through which both red light and temperature may regulate plant development [39].

Light regulation of the circadian clock

As discussed in Chapter 1, the key red light photoreceptor, phyB, seems to have two effects on circadian rhythms. First, when de-etiolated plants are exposed to red light for the first time, acute 'bursts' of transcription of *CCA1*, *LHY*, *GI*, and *PRR9*, as well as the clock regulated *CAB2* are observed [23, 142–145]. These acute 'bursts' are partly under the control of phyB and are proposed to occur due to histone modifications, whereby acetylation of histone markers H3K9 and H3K27 near transcription start sites correlates with the 'turning on' of gene expression [143, 145]. Second, the circadian rhythms following the acute 'bursts' of gene expression have altered amplitudes compared to WT rhythms. In red light, *phyB* loss-of-function mutants have lower amplitude rhythms with a larger effect under constant red light conditions compared to red 12L:12D diurnal cycles [143]. Furthermore, period lengths of circadian rhythms from *phyB* loss-of-function transgenic plants are longer than WT plants when grown in fluences of constant red light greater than ~1μmol m^{-2}

s^{-1} [9, 10]. In order to simplify the problem investigated in this chapter, I shall concentrate on finding a mechanism that can explain the impact of *phyB* loss-of-function mutants on amplitude changes of red light diurnal rhythms and the long period phenotype seen in constant red light.

Interestingly, nearly all of the key circadian clock components have been linked to the phyB-dependent light-regulation of the circadian clock. The key morning-expressed protein, CCA1, has been shown to interact with the light-stable HY5 protein [211, 218]. HY5 has been proposed to act downstream of phytochromes in light signalling, interacting with PIF1/3 proteins and regulating the light-induced acute responses of gene expression [130, 219]. As discussed in Chapter 1, the PIFs are key members of red light signalling through their interaction with phyB [132, 137, 138]. Similarly to HY5, the PRR protein TOC1 has been shown to interact with PIF proteins through its C-terminal domain [216, 217]. Further PRRs have also been implicated in red light signalling. Notably, the *prr7* loss-of-function and *PRR5-ox* mutations are unable to suppress the phenotypes of *phyB* loss-of-function transgenic plants supporting a mechanism where PRRs act downstream of phyB [178, 220]. Furthermore, transgenic plants that overexpress *PRR5* mRNA do not respond to changes in red and far-red light, in a similar manner to *phyB* mutants [221]. In addition, red light leads to the fast degradation of PRR5 protein compared to other light conditions [107, 222]. As well as the day-time circadian clock components, the important evening gene ELF3 has been shown to directly interact with phyB *in vitro*, but the two genes may act independently in the regulation of hypocotyl elongation and flowering time [214, 215]. Interestingly, loss-of-function *elf4* and *gi* mutants have been shown to alter red light-signalling pathways and alter the circadian clock. Since ELF3 interacts with ELF4 and GI, as well as phyB, then ELF4 and GI may modulate red light signalling downstream of phyB-ELF3 interactions [25, 34, 223–225]. Thus, the published literature seems to suggest the existence of many mechanisms through which the circadian clock is regulated by phyB.

Temperature and phyB

As discussed in Chapter 1, temperature effects on chemical reactions are generally thought to be approximated by the Arrhenius equation, such that reaction rates are increased by a factor of 2 as temperature rises by 10°C (this is known as the Q_{10} value). phyB has been implicated in temperature signalling with observations that warm temperature promotion of flowering is enhanced in *phyB* loss-of-function mutants [149]. Similarly, the difference in hypocotyl lengths of *phyB* loss-of-function plants grown at 22°C and 28°C is greater than that of WT plants, suggesting that phyB partially suppresses temperature-mediated elongation [161]. Interestingly, recent analysis shows that *phyB* mutant plants have a reduced sensitivity to changing red light photoperiods and may not sense diurnal changes in temperature [146, 175]. Thus, it seems reasonable that phyB plays a role in mediating temperature, as well as red light, signalling in plants. However, whether phyB plays a role in the temperature regulation of the circadian clock is unknown. One of the key properties of circadian clocks is that they are temperature compensated such that the period of the clock is maintained at ~24hrs across a range of ambient temperatures [6–8]. In this chapter I will present data looking at how phyB alters the red light diurnal rhythms of clock components. By manipulating the existing circadian clock model to qualitatively match the data, I will hypothesise the mechanisms through which phyB plays a role in regulating the circadian clock at 22°C. Finally, as phyB has been shown to play a role in temperature signalling, data will be presented showing the effects of removing phyB function from the circadian clock at 27°C. A mechanism through which phyB regulates plant development across temperature ranges will then be discussed.

Methods

Experimental Methods

Experimental procedures were conducted by Dr. Julia Foreman from Dr. Karen Halliday's group (University of Edinburgh, UK). BioDare's FFT-NLLS analysis was

conducted by Dr. Tomasz Zielinski from Prof. Andrew Millar's group (University of Edinburgh, UK).

- *Growth Conditions*

Seeds were surface sterilised then 30-40 seedlings were sown on 55 mm diameter plates containing half-strength MS media (Melford, Ipswich, UK) , pH 5.8, and 1.2% agar without added sucrose. WT and *phyB-9* seeds in the Col-4 (Columbia) accession of *Arabidopsis thaliana* were stratified at 4°C for 3 days and then grown for 7 days in 12L:12D white-light cycles (80 μmol m^{-2} s^{-1} from cool white fluorescent tubes) and then transferred to 12L:12D red-light (40 μmol m^{-2} s^{-1}) at 22°C and 27°C. For chromatin immunoprecipitation experiments, seedlings were harvested on the 6^{th} day after transfer to red L:D cycles at ZT8, 12, 16 and 20 (ZT = zeitgeber time, ZT0 = dawn). For RNA analysis, seedlings were harvested on the 14^{th} day after transfer to red L:D cycles from triplicate samples at ZT0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22 and 24 into RNA*later* solution (Sigma-Aldrich, Gillingham, UK). The plants were left overnight at 4°C in the RNA*later* solution to allow full penetration into the tissue [23].

- *RNA Extraction*

RNA was extracted from the plant tissue using the Illustra RNAspin 96 RNA isolation kit (GE Healthcare, Chalfont St. Giles, UK) manually, as described [210]. Purified total RNA (1 μg) was reverse transcribed into cDNA using SuperScript VILO cDNA synthesis kit with oligo dT primers (Invitrogen/Life Technologies, Paisley, UK) according to the manufacturer's instructions. cDNA was diluted 1/10 and 1 μl used for subsequent qRT-PCR.

- *Gene expression analysis*

qPCR reactions were set up using a liquid handling robot (freedom Evo, TECAN, Reading, UK) and run in a Lightcycler 480 system (Roche, Burgess Hill, UK) using LightCycler 480 SYBR green master mix (Roche, Burgess Hill, UK). Data was

analysed with Roche Lightcycler 480SW 1.5 using relative quantification based on the 2nd derivative maximum method. Each cDNA sample was assayed in triplicate. The primers used for in this experiment can be found in Appendix C. Transcript levels were normalized to *ACT7* expression [192].

Computational Methods

The 'Pokhilko2012' circadian clock model (see Chapter 1) was simulated using the ode15s differential equation solver on Matlab R2008b (Mathworks, Cambridge, UK; [39]). This clock model also includes induction of *TOC1* mRNA by abscisic acid (ABA). However, removing this term does not alter the dynamics of the clock system.

The SASSy (Sensitivity Analysis Software for Systems) Matlab toolbox was used to find the parameters that are able to alter the amplitude of diurnal oscillations without changing the phase (Warwick University, UK; [226]). This analysis outputs how the phase, P, and amplitude, A, of simulations change due to perturbations in parameter values, k_j. Relevant parameters were selected based on minor changes in the phase ($abs(k_j dP/dk_j) < 0.01$) with no change or a decrease in amplitude ($k_j dA/dk_j < 0.005$) for simulated rhythms of all mRNA components in the 'Pokhilko2012' clock model.

Period analysis of simulations was conducted using the FFT-NLLS (Fast Fourier Transform – Non Linear Least Squares) analysis on the BioDare data repository (www.biodare.ed.ac.uk; [227]).

Results

phyB regulates the circadian clock in red light diurnal cycles

As discussed above, previous studies have shown that phyB regulates the period of the circadian clock in constant red light, such that rhythms have a longer period in *phyB* loss-of-function plants [9, 10]. To further investigate the mechanism through

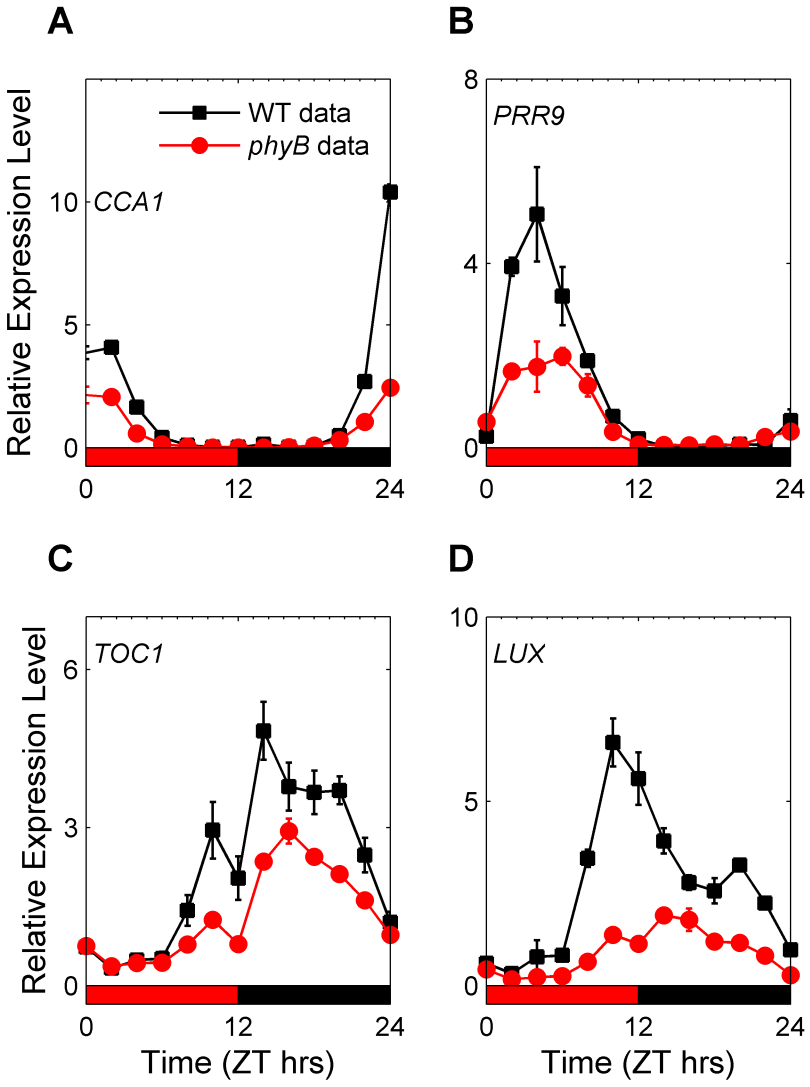


Figure 5.1: Regulation of diurnal rhythms of circadian clock components by phyB at 22°C. Transcription rhythms of (A) *CCA1*, (B) *PRR9*, (C) *TOC1*, and (D) *LUX* mRNA from red light 12L:12D cycles in WT and *phyB* loss-of-function mutant lines. Error bars represent standard error.

which phyB regulates components of the circadian clock, mRNA rhythms were obtained in red light 12L:12D diurnal cycles at 22°C (J. Foreman, University of Edinburgh, unpublished; Figure 5.1 and Figure C.1, Appendix C). As seen in Figure 5.1, the amplitudes of mRNA rhythms are lower in *phyB* loss-of-function mutants compared to WT. Thus, to hypothesise potential phyB-dependent mechanisms within the circadian clock mechanism, I manipulated the 'Pokhilko2012' circadian clock model (that was constructed to describe plants grown at ~22°C under predominantly white light) to find a mechanism that would provide low amplitude rhythms of

components in diurnal cycles and long period phenotypes in constant light [39]. As stated in Chapter 1, this clock model incorporates the EC and transcriptional inhibition of *CCA1/LHY*, *PRR9*, *PRR7*, *PRR5*, *GI*, *ELF4* and *LUX* transcription by TOC1 protein.

Parameter scanning highlights two potential mechanisms that provide simulations with a qualitative match to phenotypes of phyB loss-of-function mutants

To determine the potential phyB-dependent mechanisms within the circadian clock, two methods were used to find simulations that would decrease the amplitude of the majority of clock components in 12L:12D diurnal cycles. If these parameter changes also led to a long period phenotype in constant conditions then they can be postulated as being part of a phyB-dependent mechanism. A similar methodology was recently used to highlight how temperature- and blue light-signalling regulates the circadian clock [8]. However, in this example, only parameters that were already regulated by light were analysed. To determine whether it is possible for light to regulate the circadian machinery through a novel mechanism not already included in the mathematical clock model, I analysed the change in mRNA dynamics for any parameter change.

The first method was to manually alter the parameters of the circadian clock model between 0% and 100% (in 25% increments) of their published values (Figure 5.2a). The parameters that describe the acute light response of *CCA1/LHY*, *PRR9*, and *GI* transcription were automatically set to half of their published values in line with the observed acute response of *CAB2* luciferase constructs that is approximately halved in *phyB* loss-of-function transgenic plants [143]. In doing this, increasing the concentration of the EC, by decreasing the complexes rate of degradation, led to the suppressed amplitude rhythms of circadian clock components seen in *phyB* loss-of-function mutants in diurnal cycles (Figure 5.2a). In the current circadian clock model EC degradation is indirectly regulated by light. However, EC dynamics are strongly governed by processes that are known to be biologically light regulated. One such mechanism is the degradation of ELF3 by COP1 [34]. COP1 has been shown to

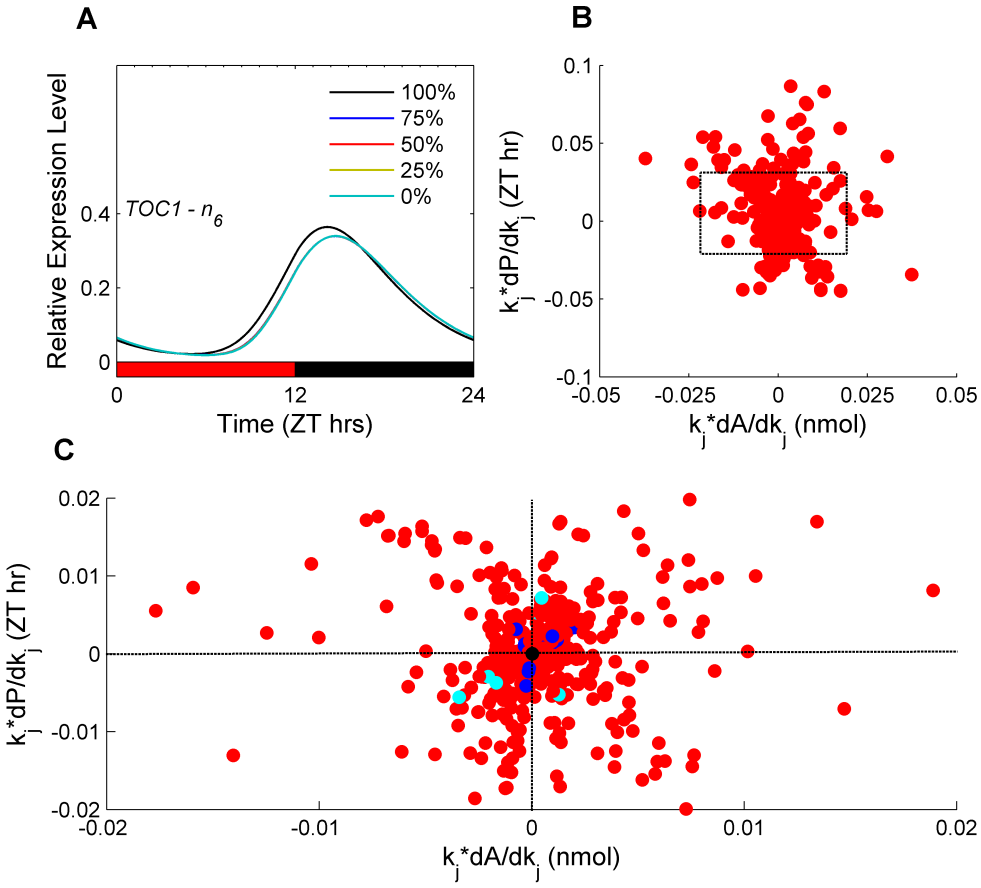


Figure 5.2: Parameter scanning to find potential phyB-dependent mechanisms in the circadian clock. (A) Simulations of mRNA components of clock model, e.g. *TOC1*, were examined when individual parameters, e.g. n_6 (light-dependent $COP1_d$ to $COP1_n$ transition), of the clock model were perturbed from original value (black line) in 25% increments (75% = blue line; 50% = red line; 25% = yellow line; 0% = cyan line). (B) Output of SASSy toolbox showing how changes in individual parameters, k_j, alters amplitude, A, and phase, P, of individual components of the circadian clock model (red dots represent changes in variables in response to parameter changes, see Computational Methods). (C) As in (B) for squared region. Black dots = n_6 (light-dependent $COP1_d$-$COP1_n$ switching); cyan dots = p_4 (TOC1 translation rate); blue dots = m_6, m_7, m_8 (TOC1 protein degradation).

translocate from the cytosol into the nucleus of plant cells during periods of darkness [228]. The action of COP1 has been proposed to be light dependent due to interactions with CULLIN 4 (CUL4), a component of an E3 ligase, and SUPPRESSOR OF PHYA (SPA) proteins [211, 229, 230]. The proposed mechanism is that during the day CUL4 is active to degrade target proteins, whilst at night COP1 takes on this function to degrade the same targets (termed as $COP1_d$ and $COP1_n$ in the 'Pokhilko2012' model, respectively; [12, 39, 211]). Therefore, the dynamics of

ELF3 protein levels, and the EC, are regulated by the light-dependent $COP1_d$-$COP1_n$ switch [34, 39]. The example highlighted in Figure 5.2a shows that by decreasing the rate of light-dependent switching between $COP1_d$ and $COP1_n$ action, leading to decreased $COP1_d$ activity and increased $COP1_n$ function, results in increased concentrations of the EC and a suppression of *TOC1* mRNA amplitude.

To support this observation, a more theoretical approach was taken with the SASSy software [226]. Using this toolbox, it is possible to find parameters that, when altered, lead to changes in amplitude without changing the phase of diurnal rhythms (see Computational Methods). Performing this analysis found three clusters of parameters that, when altered, suppress the amplitude of clock components without altering the phase (Figure 5.2b). This analysis indicated that parameters related to EC dynamics (including the $COP1_d$-$COP1_n$ switch as shown in Figure 5.2c), GI-ZTL interactions and TOC1 protein levels could alter amplitude without greatly changing phase. As GI-ZTL dynamics are blue light-dependent, these parameters will be ignored in the following analysis [33]. Figure 5.2c shows how changing the parameter values for the $COP1_d$-$COP1_n$ switch (n_6, black dots), TOC1 translation rate (p_4, cyan dots) and TOC1 degradation (m_6, m_7, m_8, blue dots) lead to phase and amplitude changes in the mRNA components of the circadian clock model. Thus, increasing the rate of $COP1_d$-$COP1_n$ switching does not greatly alter the phase or amplitude of clock components (as seen in Figure 5.2a), whilst increasing the TOC1 translation rate or protein degradation rate can lead to decreased or increased amplitude of clock components, respectively, with a smaller effect on phase altering phase. Thus, this analysis provides the hypothesis that PRR genes, like TOC1, play a role in phyB signalling [217, 220, 221]. Interestingly, PRR5 protein has been shown to be degraded at a similar rate in red light as in darkness [107, 222]. Dark-dependent degradation of PRR5 and TOC1 has been shown to be ZTL-dependent, however the mechanism that leads to red light-mediated degradation of PRR5 is yet to be discovered [32, 107]. By increasing TOC1 translation rate and decreasing acute light activation of *CCA1/LHY*, *PRR9* and *GI* transcription, the 'Pokhilko2012' model is able to capture the low amplitude diurnal rhythms of clock components (Figure 5.3a-d and Figure C.2, Appendix C). Furthermore, these parameter changes result in

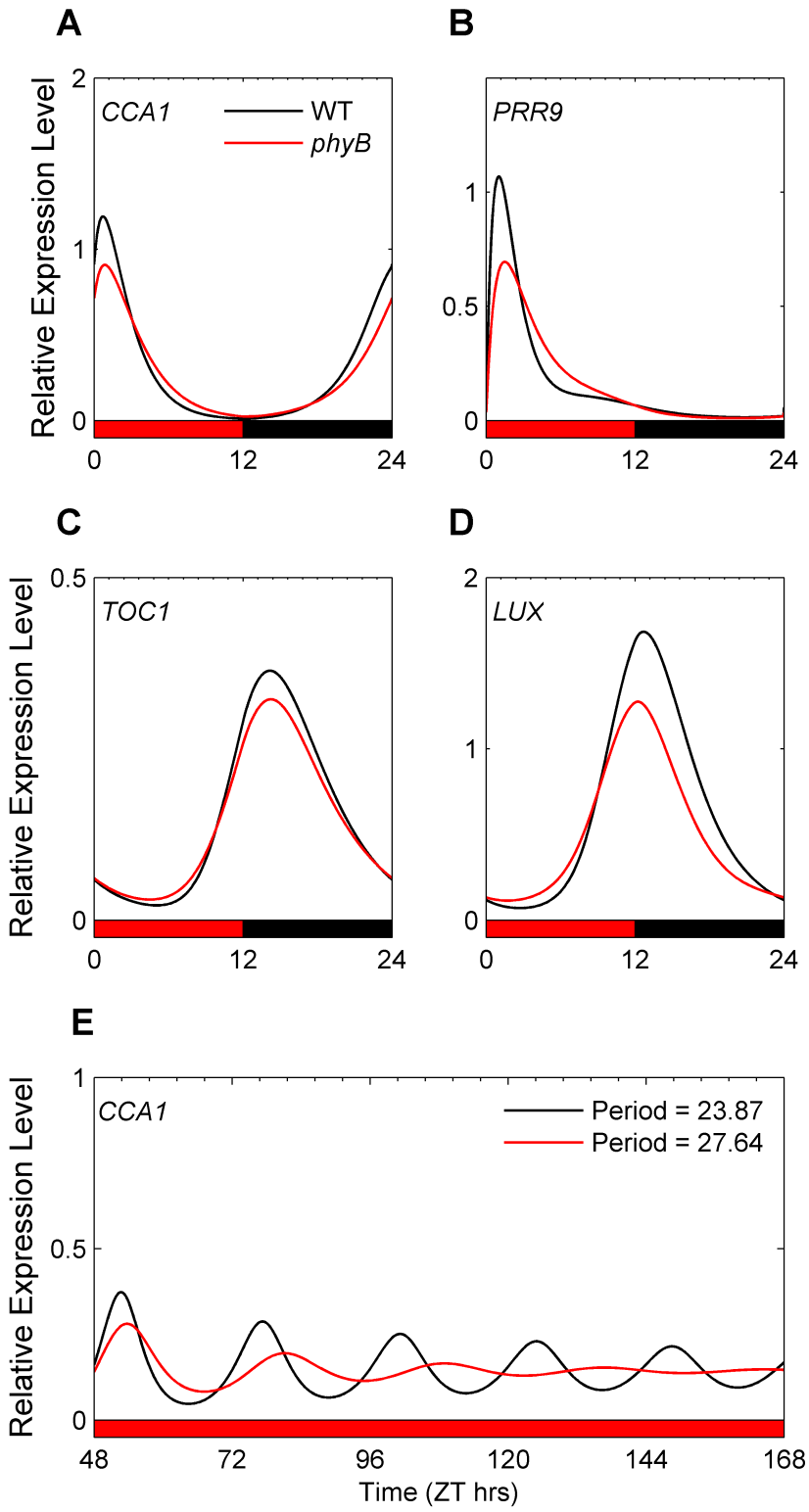


Figure 5.3: Simulations of *phyB* loss-of-function mutant through TOC1 overexpression at 22°C. Simulations of red light 12L:12D diurnal cycles of

(continued from above) cycles of (A) *CCA1*, (B) *PRR9*, (C) *TOC1*, and (D) *LUX* mRNA in WT (black line) and *phyB* loss-of-function mutant (red line). (E) Simulation of *CCA1* mRNA in constant red light conditions. Period lengths calculated by FFT-NLLS.

---

simulations that mimic the long period phenotype of *phyB* loss-of-function transgenic plants (Figure 5.3e). However, the ~4hr period lengthening of *phyB* simulations is approximately double what has been observed experimentally [9, 10]. The altered model parameter values for these simulations are given in Table 5.I. Thus, from the analysis of the circadian clock model presented here, one can hypothesise that phyB-signalling plays a role in the red light-mediated degradation of TOC1 protein such that TOC1 protein levels are overexpressed in the absence of phyB.

phyB maintains a role in the circadian clock at 27°C

To look at whether phyB plays a role in the temperature regulation of the circadian clock machinery, transcription profiles of clock components were obtained at 27°C (J. Foreman, University of Edinburgh, unpublished). Previous reports have shown that phyB inhibits temperature responses such that when phyB function is removed,

Table 5.I: Changes in parameter values to describe the *phyB* loss-of-function mutant by TOC1 protein overexpression.

k_j	**Parameter Function**	**WT value**	***phyB* value**
q_1	Acute light activator of *CCA1/LHY* transcription	1.2	0.6
q_2	Acute light activator of *GI* transcription	1.56	0.78
q_3	Acute light activator of *PRR9* transcription	3	1.5
m_{11}	Degradation of acute light activator	1	0.5
p_4	TOC1 translation rate	0.5	1

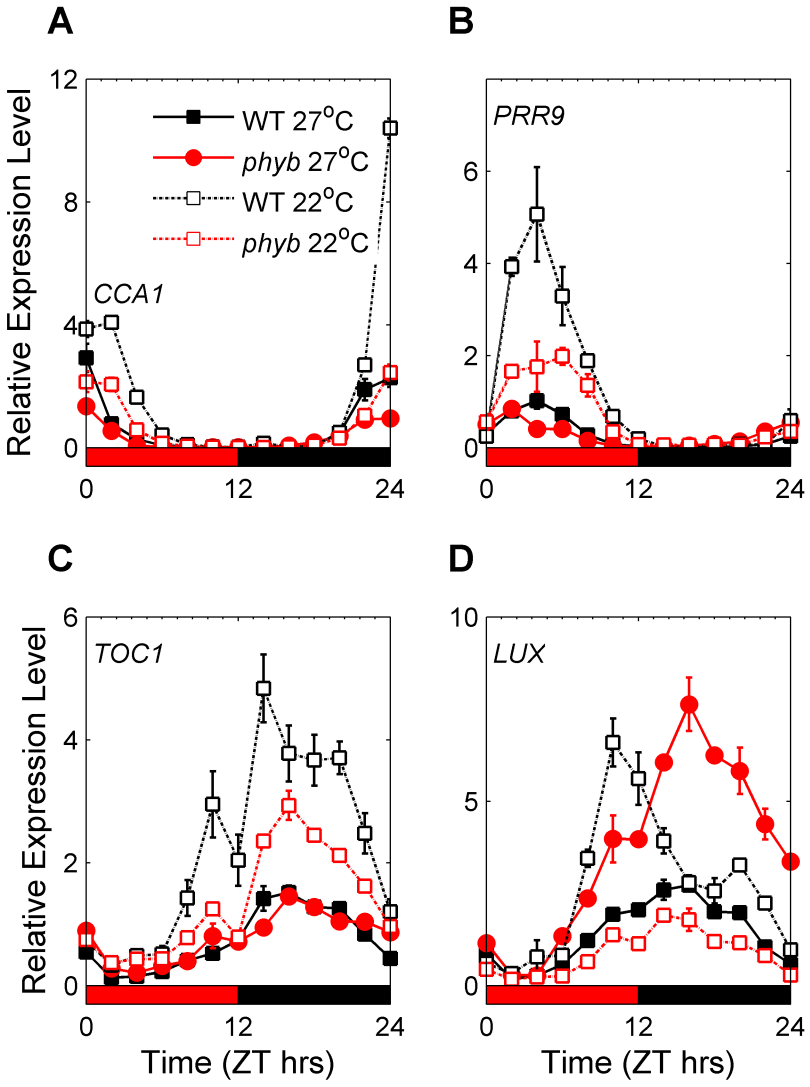


Figure 5.4: Regulation of the circadian clock by phyB at 27°C. As in Figure 5.1 for red light 12L:12D diurnal cycles at 27°C. Data from Figure 5.1 shown with empty symbols and dashed lines. Error bars represent standard error.

temperature effects are enhanced [149, 161]. Comparing the expression of clock components in WT rhythms at 22°C and 27°C showed that the amplitude of gene expression was decreased in the warm, in a similar manner to the removing of phyB at 22°C (Figure 5.4 and Figure C.3, Appendix C). However, the amplitude of diurnal rhythms of the clock components does not always decrease further in *phyB* loss-of-function transgenic plants grown at 27°C. In some cases, such as *CCA1* and *PRR9* mRNA, the amplitude does decrease in *phyB* mutants compared to WT at 27°C (Figure 5.4a and b). Conversely, the removal of phyB results in very minor changes

to the rhythm of *TOC1* mRNA but, remarkably, leads to a large increase in *LUX* transcription during 27°C nights (Figure 5.4c and d). Interestingly, though, examples such as the transcription of *PRR7* and *GI* may show that transcription rhythms controlled by diurnal cycles are altered by the removal of phyB such that they maintain relatively constant levels throughout a whole day (Figure C.3a and c, Appendix C).

To explain these results, two hypotheses for how both light and temperature effect the clock can be postulated. First, the regulation of the circadian clock by phyB changes across temperatures, such that at 22°C phyB regulates a single component (e.g. TOC1 or the EC) but, at 27°C, phyB regulates multiple components (possibly TOC1 and the EC) of the clock machinery. This system would lead to decreased amplitude rhythms of clock components at 22°C in *phyB* mutants and may also go some way to explaining the expression patterns observed at 27°C where the loss of phyB has multiple effects on transcriptional rhythms. Conversely, a counter hypothesis is that the effects of temperature and phyB can be separated in the circadian clock mechanism. In this system, the loss of phyB and the increase in temperature both alter the clock in the same manner (i.e. decreasing the amplitude of mRNA rhythms) albeit through independent mechanisms. Thus, when *phyB* loss-of-function transgenic plants are grown at 27°C more than one clock function is modified; one alteration of the clock would be due to the removal of phyB activity whilst the other is due to increased growth temperatures. The resulting rhythms of *phyB* mutants grown at 27°C would then appear to be significantly different to the WT rhythms as seen in Figure 5.4. Therefore, future experimental efforts should concentrate on determining how the clock machinery is altered across temperatures and whether these changes are phyB-dependent.

Discussion

In this chapter I have used the most recent published circadian clock model, 'Pokhilko2012', to represent WT conditions at 22°C in red light and hypothesise a mechanism through which phyB and temperature may regulate the circadian clock

[39]. This forms part of on-going work with collaborators from Dr. Karen Halliday's group (University of Edinburgh, UK) whom are experimentally testing the hypothesis derived in this chapter.

The model-derived hypothesis states that phyB can regulate the circadian clock through either changing EC (through COP1) or TOC1 protein dynamics, or potentially both together. This conclusion was reached by testing which parameters are able to provide simulations that qualitatively match the low amplitude diurnal rhythms of clock components and long period phenotype in constant red light. Simulations of overexpressed TOC1 protein provided both these requirements (compare Figure 5.1 with 5.3; [9, 10]). This suggests that phyB may negatively regulate either TOC1 levels or activity. Interestingly, a further member of the PRR protein family, PRR5, has been shown to be degraded faster in red light compared to white or blue light [107, 222]. Thus it seems plausible that PRR proteins play an important role in red light signalling, as has been observed with hypocotyl and flowering phenotypes [220]. phyB has previously been shown to degrade PIF proteins through a mechanism that involves phosphorylation and ubiquitination of the PIF proteins [132, 137, 138]. Thus, it may be possible for PRR proteins, such as TOC1 and PRR5, to be degraded through a similar system. Alternatively, TOC1 and PIF3 proteins have been shown to interact through yeast 2-hybrid experiments, suggesting that an interaction between TOC1 and phyB may occur through the PIF proteins [216, 217].

Interestingly, the effects of increased temperature (to 27°C) have the same effect on diurnal rhythms of circadian clock components to the removal of phyB at 22°C, resulting in decreased amplitudes of mRNA expression (Figure 5.4 and Figure C.3, Appendix C). As discussed above, the analysis presented here has found two mechanisms through which low amplitude diurnal rhythms could be achieved using the 'Pokhilko2012' circadian clock model [39]. Thus, it is possible that the same single mechanism could mediate both phyB-signalling into the circadian clock whilst controlling the clocks response to a 5°C temperature increase. This would also result in the period length of the circadian clock to be greater than 24 hours in both *phyB*

loss-of-function mutants at 22°C and WT plants at 27°C grown in constant red light as has been experimentally observed [9, 10]. However, having phyB and temperature working through the same single mechanism would not be able to account for the diurnal rhythms of *LUX*, *PRR7* and *GI* mRNA from *phyB* loss-of-function plants grown in 27°C conditions (Figure 5.4d and Figure C.3a and c, Appendix C). These rhythms do not have synergistically lower amplitude as a result of increased temperature and the removal of phyB function. Thus, it appears that more than one element of the circadian clock machinery is altered when temperature is increased and phyB function is removed at the same time. Such a separation of phyB and temperature mechanisms would allow for the entrainment of the circadian clock by H:C cycles in the absence of light cues (i.e. in darkness) [164]. It would be of interest to determine whether the two mechanisms required for temperature and phyB responses, together, feature TOC1 protein and the EC since overexpression of one of these components results in the low amplitude, long period rhythms in the circadian clock model that were analysed in this chapter.

A further question to explore is how plant phenotypes, such as the timing of flowering and hypocotyl elongation, at warm temperatures are enhanced by the loss of phyB function [149, 161]. Given that the data presented in this chapter shows that the majority of circadian clock components are not as severely perturbed in *phyB* loss-of-function plants compared to WT at 27°C, one possibility is that phyB acts downstream of the circadian clock, thus forming a feed-forward network. The key regulators of hypocotyl elongation and flowering in Arabidopsis, PIF4 & 5 and CO respectively, have been shown to be strongly regulated by phyB mediated red light-signals [86, 138]. Furthermore, PIF4 protein levels have been shown to be regulated by temperature [125, 162]. In Chapter 4, with the use of mathematical modelling I hypothesised a mechanism whereby the PIFs and CO interact to regulate flowering at 22°C and 27°C. Interestingly, the CCT domain of TOC1 proteins that PIF3 interacts with is present in the CO protein structure [216, 217]. Thus, phyB may be able to suppress temperature responses in plants through mediating temperature-responsive PIF interactions with TOC1, CO and further partners.

In the next chapter I shall bring together the conclusions made throughout this thesis and speculate as to the future directions of research that are required to get a better understanding of the environmental regulation of plant development.

Chapter 6: Summary & Conclusions

Since plants are sessile organisms, crop development is strongly linked to the local environment. However, as climate change continues and environmental fluctuations become harder to predict, it is important for scientists (and society at large) to understand how the altered growth conditions of crops will affect yields and biofuel feedstock production. By determining how plants are able to cope with fluctuating environments, crop management may be further improved to prevent the loss of yield in adverse seasonal conditions. In this thesis, I have developed mathematical models that aim to further our understanding of the molecular processes that link the local growth environment of plants with their development. By using the model plant *Arabidopsis thaliana* as a case-study, I have linked flowering with day-length, light- and temperature-signalling – three important environmental cues sensed by plants. Furthermore, by looking at how the circadian clock (the principle day-length measuring mechanism within plants) is regulated by changes in red light- and temperature-signals, the integration of environmental signals is starting to be elucidated by using systems biology approaches.

In Chapter 2, I used model selection techniques to construct a system that described the circadian regulation of *PIF4* and *PIF5* transcription across multiple photoperiods. The use of model selection techniques helps determine the simplest model that accurately describe experimentally obtained mRNA rhythms. This analysis showed that the best supported model of circadian regulated *PIF* transcription requires rhythmic signals from multiple clock components to control the *PIF* genes. However, this model was built to describe rhythms of *PIF* transcription in WT plants under different photoperiods, not in transgenic mis-expression lines or mutants. Yet, in Chapter 4, models built to describe the circadian regulation of *FKF1* and *CDF1* transcription across different photoperiods and in different transgenic lines also required at least two circadian clock regulators. Thus, it seems from the models developed here, which are based on the currently available data, that regulation of

transcription by the circadian clock is a dynamic process that often features multiple components of the circadian clock.

In Chapters 3 & 4, I went on to link the circadian regulation of flowering with blue light-, red light- and temperature-signalling within mathematical models. These two models highlighted four important features of plant flowering in response to environmental signals:

1. components of the circadian clock, particularly CCA1, LHY and ELF3, may have specific roles in flowering independent of the transcriptional regulation of *FKF1* and *CDF1* by mediating levels of CDF1 protein;
2. of the two roles that the blue light-regulated protein FKF1 plays in the flowering system, photoperiodic regulation of *FT* transcription is relatively more sensitive to FKF1 stabilisation of CO than FKF1 mediated degradation of CDF1;
3. due to the similar rhythms of *CO* transcription seen in WT and *pif4;5* loss-of-function mutations, it seems that red light-signalling mediated by PIF4 and PIF5 regulates LD flowering (and *FT* transcription) through a CO-PIF protein interaction, and;
4. temperature-sensitivity of flowering is dependent on photoperiod such that, in LDs, temperature sensitivity is mediated through the photoperiodic flowering pathway, notably temperature-sensitive CO protein stability, whereas in SDs PIF4 is required for the acceleration of flowering in warm temperatures (~27°C).

Thus the regulation of CO protein by external signals seems to be critical for the correct regulation of flowering across fluctuating environmental conditions such that CO stability is increased in warm ambient temperatures (~27°C) and in the absence of red light signals [86, 153]. Yet, it is unknown whether an interaction between CO and PIF4, a similarly light- and temperature-regulated protein, is required for the response of CO protein levels to external signals [137, 162]. Interestingly, PIF3 has already been shown to interact with TOC1 protein through the CCT domain that is shared with the CO protein structure [216, 217]. Furthermore, both CO and PIF4

have been shown to be localised to the vascular tissues of leaves [174, 231]. However, the protein abundance of CO *in planta* is low such that *CO* overexpression lines need to be used to detect the diurnal rhythm of CO protein and interactions between CO and other proteins [86, 153]. Thus, whilst a PIF4-CO interaction can be tested using a coimmunoprecipitation system (as used in [153] to determine a CO-FKF1 interaction), the overexpression of *CO* in plants may lead to altered protein complex dynamics such that CO sequesters the PIF4 protein. If CO senses environmental signals through PIF4, sequestration could result in CO protein levels being more sensitive to environmental cues in *CO* overexpression lines than in WT plants.

Finally, in Chapter 5, I examined how red light- and temperature-signals regulate the circadian clock mechanism, thereby integrating these signals directly with the sensing of photoperiod in Arabidopsis. By using a pre-existing mathematical model of the circadian clock, I hypothesised a mechanism through which red light signals may alter the circadian clock. The resulting analysis highlighted the potential for red light, through phyB, to regulate TOC1 protein dynamics. It is interesting to note that, like CO protein, TOC1 contains a CCT domain that is important for DNA-binding activities [232, 233]. Furthermore, the CCT containing C-terminus of TOC1 has been shown to interact with PIF proteins that are important in red light- and temperature-signalling [216, 217]. Thus, it remains to be seen whether the CCT domain is important for the integration of multiple environmental signals in the circadian clock and flowering systems.

Whilst the modelling approach employed throughout this thesis has provided a number of experimentally testable hypotheses that have been validated, there are limitations to what the models can describe due to assumptions that have been made during model construction. For example, as the flowering model presented in Chapter 4 is built as an output system of the 'Pokhilko2011' circadian clock model then any limitations to the clock model will be maintained in the flowering system. For example, the clock models description of GI-ZTL, GI-ELF3 and ELF3-COP1 protein interactions has a direct impact on the flowering system. In the

'Pokhilko2011' clock model the nuclear pool of GI protein is negatively regulated by ELF3- and COP1-mediated sequestration and degradation, whilst the cytoplasmic pool is sequestered into GI-ZTL complexes [12]. Thus, a large proportion of available GI in the system is already used before it can form complexes with FKF1 and CDF1 in the flowering system. Furthermore, FKF1 has been shown to have only a minor effect on the period of circadian rhythms suggesting that it may not interact with a large enough pool of GI to alter function of GI in the circadian clock [109]. Therefore, rather than directly model the protein dynamics of FKF1-GI complex formation, I assumed that FKF1 and GI interact to degrade CDF1 protein without altering the pools of GI that are required for the circadian clock model to function correctly. The result of this assumption is that the short period lengthening of the circadian clock observed in *fkf1* loss-of-function transgenic plants (possibly due to the availability of a larger pool of free GI) cannot be replicated here [109].

A further limitation to the model presented in Chapter 4 is that it does not take into account any regulators of *FT* transcription that act independently of CO protein. Thus, whilst experimentally PIF4 has been shown to regulate flowering without the presence of CO protein in SDs, the presented model would be unable to describe this effect since PIF4 and CO proteins act together to control *FT* expression levels in LDs [174]. To disseminate the two roles of PIF4, one that is CO-dependent and the other that is CO-independent, time-series data of *FT* mRNA would need to be produced in SDs in WT and transgenic plants (such as *pif4;5* and *CO-ox;pif4;5* or *co;pif4;5*). From this data, parameters could be constrained such that CO-independent PIF4 activation of *FT* transcription can be included in the flowering model. This may then allow for the inclusion of further components that regulate *FT* transcription under specific conditions, such as HOS1 that controls flowering in the cool (~16°C) through a partially CO-independent mechanism (see Chapter 1; [171]).

From the conclusions presented above, there are a number of directions in which future research of this area could elucidate new information and develop on the ideas presented in this thesis. These directions cross multiple scales from examining the molecular level in more depth, to coordinating temporal regulation across spatial

domains, before finally looking at how molecular mechanisms could integrate into whole plant models and crop studies. Each of these would provide us with a better understanding of how the environment and the circadian clock are able to correctly regulate plant development.

First, it seems crucial to obtain a deeper understanding of protein regulation across different day-length, light- and temperature-regimes. For the development of the flowering models, a number of quantified protein time-series data were used to obtain a better understanding of, for example, the regulation of CDF1 protein levels by the GI-FKF1 complex (see Chapter 4). However, in the construction of the circadian clock models, not many quantified protein datasets have been used to constrain simulated protein dynamics. Thus, it is hard to know how/whether a circadian clock protein is able to interact or regulate the stability of other clock proteins. Given the hypothesis presented in Chapter 4 that CCA1, LHY and ELF3, as well as the already published role of GI, are able to regulate CDF1 protein levels, this underlines the conclusion that the roles of circadian clock proteins are not just restricted to the circadian clock mechanism but may also have specific roles in communicating timing information to plant developmental pathways. Whilst transcriptional control of output targets by the circadian clock has been previously discovered and described throughout this thesis, the suggestion that clock components may also regulate developmental pathways post-translationally is a novel hypothesis to be tested in future studies [20, 25, 31].

Additionally, as more becomes known about the dynamics of circadian clock proteins, it would be of interest to examine the mechanisms through which the circadian clock regulates transcription. Whilst current knowledge suggests that circadian clock proteins regulate target genes by binding to promoter motifs at specific times of the day (see Chapter 2), recent analysis of the mammalian circadian clock suggests a more stochastic view of circadian regulation [53, 54]. What these studies highlighted was that rhythmic gene expression does not occur through a continuum of circadian regulation where, for example, CCA1 & LHY regulate target gene expression in the morning, followed by PRRs during the day and the EC at

night, but that transcriptional regulation occurs in bursts. Interestingly, some individual promoter motifs are targeted by multiple circadian clock proteins in Arabidopsis. Notably, promoters targeted by CCA1, LHY and TOC1 have a significant number of EE-like motifs suggesting that the EE motif may be important for both CCA1/LHY- and TOC1-mediated transcriptional regulation [19, 30]. Therefore, the recycling of one clock protein, e.g. CCA1 or LHY, on the promoter may allow for a second protein, e.g. TOC1, to regulate transcription of the target gene in fluctuating 'bursts'. Conversely, the convergence of multiple circadian regulators on the promoters of target genes may enhance the stability of proteins by preventing degradation by the proteasome on the promoter. The proteasome is required for the fast turnover of proteins on target promoters resulting in stochastic 'bursts' of circadian regulated transcription in mammalian systems [54]. One system that may help elucidate the dynamics of transcriptional regulation *in planta* could be the photoperiodic flowering network outlined in Chapters 3 & 4. In this system, CO, FKF1, GI and CDF1 have been shown to regulate each others stability whilst also occupying similar locations on the promoter of *FT*. Thus, if the continual turnover of these proteins on the *FT* promoter is required for the correct regulation of *FT* transcription, then it may point to a more general mechanism of transcriptional control throughout the whole plant.

Whilst the two experimental directions outlined above require technically demanding experiments to obtain dynamic information about specific proteins in the circadian and flowering systems, the models presented in this thesis can also be used to understand how the circadian clock and flowering integrates with other developmental pathways regulating the whole plant. Recent work has been conducted to elucidate how FT protein dynamics regulate flowering [75, 234]. As well as the temporal regulation of *FT* transcription, spatial transport of FT protein is required for the induction of flowering. Thus, it seems that *FT* is transcribed in a photoperiod-dependent manner in leaves before the protein translocates to the shoot apical meristem (SAM) where it induces flowering through transcription of *APETALA 1* (*AP1*) [75]. Mathematical analysis of this system was able to describe floral induction after FT protein had been transported to the SAM [234].

Furthermore, and potentially as a result of transcriptional regulation by FT protein, stomatal aperture (opening and closing of the stomata) has been shown to be regulated through the photoperiodic flowering pathway [235]. It would be interesting to see whether the model presented in Chapter 4 could be used as a module in larger models describing spatial FT dynamics. Importantly, given the recent analysis determining circadian rhythms on a spatial scale, the transportation of FT protein and its regulation of transcription may be affected by spatial heterogeneity of circadian clock function [236, 237].

Finally, given the importance in determining the effects of the environment on the whole plant, a recent mathematical model was able to predict the flowering of multiple accessions of Arabidopsis from field studies [175]. Using information about the local environments where the field studies were conducted, the time taken for a plant to flower was estimated with calculations based on recorded day-lengths and temperature cycles. Given that this environmental information is now encompassed in the flowering model constructed in Chapter 4, it may be possible to use the model presented in this thesis to predict flowering in field studies. Thus, the model of Chapter 4 could be used as a small part of a larger system describing multiple molecular mechanisms of Arabidopsis plants. This has implications for whole plant models and computational crop studies. In these systems, large models are constructed by combining multiple smaller models describing more specific molecular processes. This modular technique has been used recently to model the life cycle of a human pathogen, *Mycoplasma genitalium* [238]. By combining multiple whole plant models, models would be able to predict what occurs in fields of crops. Hence, by obtaining a better understanding of the molecular mechanisms regulating an individual plant, the science community is able to translate the results to other plant species and resulting crop yields through modular models. This approach might contribute to understand the effects of altered environments on the yields of particular crop genotypes.

Appendix A: Parameter values of model presented in Chapter 3

Table A.I: Parameter values of model in Chapter 3 as obtained by simulated annealing.

Parameter	Interpretation	Value	Units
n_1	CDF1-dependent *CO* transcription	4.33	nmol/hr
n_2	CDF1-independent *CO* transcription	1.365	nmol/hr
n_3	CDF1-independent *FT* transcription	0.135	nmol/hr
n_4	CDF1-dependent *FT* transcription	1.546	nmol/hr
m_1	GI/FKF1-independent CDF1 degradation	0.3344	1/hr
m_2	*CO* mRNA degradation	0.864	1/hr
m_3	Basal CO protein degradation	1.0851	1/hr
m_4	Dark-dependent CO protein degradation	38.3384	1/hr
m_5	*FT* mRNA degradation	0.243	1/hr
g_1	Michaelis-Menten coefficient for CDF1 inhibition of *CO*	0.055	nmol
g_2	Michaelis-Menten coefficient for FKF1-independent GI disruption of CDF1 activity	0.00005	nmol
g_3	Michaelis-Menten coefficient for activation of *CO*	2	nmol
g_4	Michaelis-Menten coefficient for FKF1-CO stabilization	0.1	nmol
g_5	Michaelis-Menten coefficient for CDF1 inhibition of *FT*	0.165	nmol
g_6	Michaelis-Menten coefficient for CO activation of *FT*	0.276	nmol
p_1	*CDF1* mRNA translation	0.09825	1/hr
p_2	GI/FKF1-dependent CDF1 degradation	7.74706	1/hr

p_3	*CO* mRNA translation	0.56413	1/hr
p_4	CO stabilization by FKF1	4.4484	1/hr
a	Hill coefficient	2	
b	Hill coefficient	2	
c	Hill coefficient	2	
B_{CO}	Basal *CO* transcription	0.049	nmol/hr

Appendix B: Parameter values of model presented in Chapter 4

Photoperiodic Hypocotyl Elongation Model

This model was constructed by Dr. Daniel Seaton, a collaborator in Prof. Andrew Millar's group (University of Edinburgh, UK). The model equations describing photoperiodic hypocotyl elongation are:

$$\frac{dc_{PIF4}^{(m)}}{dt} = n_{10}\frac{g_{10}}{g_{10}+c_{EC}(\tau)}\left(1+n_{11}\frac{{c_L}^f}{g_{11}^f+{c_L}^f}\right)-m_9c_{PIF4}^{(m)} \tag{B.1}$$

$$\frac{dc_{PIF5}^{(m)}}{dt} = n_{12}\frac{g_{12}}{g_{12}+c_{EC}(\tau)}\left(1+n_{13}\frac{{c_L}^g}{g_{13}^g+{c_L}^g}\right)-m_{10}c_{PIF5}^{(m)} \tag{B.2}$$

$$\frac{dc_{phyB}}{dt} = p_9L(1-c_{phyB})-m_{11}c_{phyB} \tag{B.3}$$

$$\frac{dc_{PR}}{dt} = p_{10}L(1-c_{PR})-m_{12}c_{PR} \tag{B.4}$$

$$\frac{dc_{Int}}{dt} = p_{11}+p_{12}c_{PR}-m_{13}c_{Int}-p_{13}c_{Int}c_{PIF}+m_{14}c_{PIF.Int} + (m_{15}+m_{16}c_{phyB})c_{PIF.Int} \tag{B.5}$$

$$\frac{dc_{PIF}}{dt} = (p_{14}c_{PIF4}^{(m)}+p_{15}c_{PIF5}^{(m)})-(m_{15}+m_{16}c_{phyB}).c_{PIF}-p_{13}c_{Int}c_{PIF} + m_{14}c_{PIF.Int}-2p_{16}c_{PIF}^2+2m_{17}c_{PIF_{dimer}}+m_{13}c_{PIF.Int} + 2(m_{15}+m_{16}c_{phyB})c_{PIF_{dimer}} \tag{B.6}$$

$$\frac{dc_{PIF.Int}}{dt} = p_{13}c_{Int}c_{PIF}-m_{14}c_{PIF.Int}-(m_{13}+m_{15}+m_{16}c_{phyB})c_{PIF.Int} \tag{B.7}$$

$$\frac{dc_{PIF_{dimer}}}{dt} = p_{16}c_{PIF}^2 - m_{17}c_{PIF_{dimer}} - 2(m_{15} + m_{16}c_{phyB})c_{PIF_{dimer}} \quad \text{(B.8)}$$

where parameters are denoted in the same fashion as in Chapter 3 and 4. τ represents a τ-hour phase advance of EC dynamics required to obtain accurate rhythms of *PIF4* and *PIF5* mRNA across a range of photoperiods. These equations can be added to the 'Pokhilko2012' circadian clock model to describe regulation of PIF protein levels required for *FT* transcription (equation 4.7). In equation (4.7), $c_{PIF} = c_{PIF} + c_{PIF.Int} + c_{PIF_{dimer}}$. In this system, components phyB, PR and 'Int' represent the red light photoreceptor phytochrome B, a non-phyB red light photoreceptor and a number of PIF interacting proteins, respectively. Each of these groups have been shown experimentally to regulate PIF protein stability and dynamics [121, 125, 129, 131, 137, 139, 141, 239, 240]. This model was constructed by Daniel Seaton, a collaborator in the Millar group (University of Edinburgh, UK). Parameter values for this system are given in Table B.II.

Table B.I: Parameter values of model in Chapter 4 as obtained by simulated annealing.

Parameter Name	Interpretation	Parameter Value	Units
n_1	CCA1/LHY-independent *CDF1* transcription	0.2669	nmol/hr
n_2	CCA1/LHY-dependent *CDF1* transcription	2.329	nmol/hr
n_3	*FKF1* transcription	3.089	nmol/hr
n_4	NP-independent *CO* transcription	0.4122	nmol/hr
n_5	NP-dependent *CO* transcription	1.1211	nmol/hr
n_6	PIF-independent *FT* transcription	0.8012	nmol/hr
n_7	PIF-dependent *FT* transcription	0.492	nmol/hr
n_8	CDF1-independent *FT* transcription	0.0324	nmol/hr
n_9	CDF1-dependent *FT* transcription	0.371	nmol/hr
m_1	*CDF1* mRNA degradation	0.43	1/hr

m_2	Basal CDF1 protein degradation	0.341	1/hr
m_3	*FKF1* mRNA degradation	1.353	1/hr
m_4	Basal FKF1 protein degradation	283	1/hr
m_5	*CO* mRNA degradation	0.767	1/hr
m_6	COP1-dependent CO protein degradation	38.3384	1/hr
m_7	Basal CO protein degradation	1.0859	1/hr
m_8	*FT* mRNA degradation	0.243	1/hr
g_1	Michaelis-Menten coefficient for CCA1/LHY activation of *CDF1*	0.94	nmol
g_2	Michaelis-Menten coefficient for PRR inhibition of *CDF1*	0.789	nmol
g_3	Michaelis-Menten coefficient for CCA1/LHY inhibition of *FKF1*	0.28	nmol
g_4	Michaelis-Menten coefficient for EC inhibition of *FKF1*	0.006	nmol
g_5	Michaelis-Menten coefficient for CDF1 inhibition of *CO*	0.082	nmol
g_6	Michaelis-Menten coefficient for NP activation of *CO*	1.9908	nmol
g_7	Michaelis-Menten coefficient for PIF activation of *FT*	16.463	nmol
g_8	Michaelis-Menten coefficient for CDF1 inhibition of *FT*	0.0351	nmol
g_9	Michaelis-Menten coefficient for CO activation of *FT*	0.828	nmol
p_1	*CDF1* mRNA translation	0.0694	1/hr
p_2	GI-FKF1-dependent CDF1 degradation	0.7859	1/hr
p_3	GI-dependent, FKF1-independent CDF1 degradation	0.0293	1/hr
p_4	ELF3-dependent CDF1 stabilisation	1.4663	1/hr
p_5	*FKF1* mRNA translation	1.115	1/hr
p_6	GI-dependent FKF1 stabilisation	0.005	1/hr

p_7	*CO* mRNA translation	3.0883	1/hr
p_8	FKF1-dependent CO stabilisation	4.4484	1/hr
k_1	Michaelis-Menten coefficient for ELF3 stabilisation of CDF1	0.04	nmol
k_2	Michaelis-Menten coefficient for GI stabilisation of FKF1	1.801	nmol
k_3	Michaelis-Menten coefficient for FKF1 stabilisation of CO	0.001	nmol
q_1	Acute light activation of *FKF1* transcription	0.21	1/hr
B_{CO}	Basal *CO* transcription	0.0106	nmol/hr
a	Hill coefficient	2	
b	Hill coefficient	2	
c	Hill coefficient	2	
d	Hill coefficient	2	
e	Hill coefficient	2	
a_F	Scaling factor of *FT* expression	-2308.141	
d_0	Lowest number of days to flower	16.55	days
b_F	Lower bound for *FT* required to flower	0.02	nmol

Table B.II: Parameter values for hypocotyl elongation model.

Parameter Name	Interpretation	Parameter Value	Units
n_{10}	CCA1/LHY-independent *PIF4* transcription	1.827	nmol/hr
n_{11}	CCA1/LHY-dependent *PIF4* transcription	0.2156	nmol/hr
n_{12}	CCA1/LHY-independent *PIF5* transcription	1.9231	nmol/hr
n_{13}	CCA1/LHY-dependent *PIF5* transcription	1.638	nmol/hr
m_9	*PIF4* mRNA degradation	1.993	1/hr
m_{10}	*PIF5* mRNA degradation	4.6814	1/hr
m_{11}	phyB protein degradation	0.6391	1/hr
m_{12}	PR protein degradation	0.6391	1/hr
m_{13}	PIF 'Interactor' protein degradation	1.548	1/hr

m_{14}	De-stabilisation of PIF-'Interactor' protein interactions	5	1/hr
m_{15}	phyB-independent PIF protein degradation	1.5049	1/hr
m_{16}	phyB-dependent PIF protein degradation	3.7815	1/hr
m_{17}	Degradation of PIF dimer complexes	5	1/hr
g_{10}	Michaelis-Menten coefficient for EC inhibition of *PIF4*	0.0116	nmol
g_{11}	Michaelis-Menten coefficient for CCA1/LHY activation of PIF4	0.2972	nmol
g_{12}	Michaelis-Menten coefficient for EC inhibition of *PIF5*	0.1083	nmol
g_{13}	Michaelis-Menten coefficient for CCA1/LHY activation of PIF5	0.4893	nmol
p_9	phyB production	20	1/hr
p_{10}	PR production	20	1/hr
p_{11}	PIF 'Interactor' protein production	85.8422	nmol/hr
p_{12}	PR-dependent 'Interactor' protein production	60.1045	1/hr
p_{13}	PIF-'Interactor' protein complex formation	115.4979	1/hr
p_{14}	*PIF4* translation	43.6364	1/hr
p_{15}	*PIF5* translation	112.7338	1/hr
p_{16}	PIF dimer formation	49.9135	1/hr
f	Hill coefficient	2	
g	Hill coefficient	2	
τ	Phase altering of EC activity	2	hr

Appendix C: Data analysed in Chapter 5

Table C.I: List of primers used in qPCR experiments of Chapter 5.

Transcript	Primers
ACT7	5'-CAGTGTCTGGATCGGAGGAT-3' 5'-TGAACAATCGATGGACCTGA-3'
CCA1	5'-CTGTGTCTGACGAGGGTCGAA-3' 5'-ATATGTAAAACTTTGCGGCAATACCT-3'
LHY	5'-TAGCTAAGGCAAGAAAGCCA-3' 5'-GCTTCTAGAAACCTCTCATGC-3'
PRR9	5'-GATTGGTGGAATTGACAAGC-3' 5'-TCCTCAAATCTTGAGAAGGC-3'
PRR7	5'-CTTTCTCAAGGTATAATCCAGCC-3' 5'-ACAATCATATGCTGCTTCAGTC-3'
PRR5	5'-GTGTATGTTGAAAGGTGCGG-3' 5'-AGGAGCAAGTGAAGTTTGTC-3'
TOC1	5'-ATCTTCGCAGAATCCCTGTGATA-3' 5'-GCACCTAGCTTCAAGCACTTTACA-3'
GI	5'-TATTGAAGTGTCGTCTACCAG-3' 5'-GAGCTTTGGTTCATGATATCAC-3'
ELF3	5'-GGAAAGCCATTGCCAATCAA-3' 5'-ATCCGGTGATGCAGCAATAAGT-3'
ELF4	5'-CGACAATCACCAATCGAGAATG-3' 5'-AATGTTTCCGTTGAGTTCTTGAATC-3'
LUX	5'-TGCTCATCATCTTCACAAACC-3' 5'-CTTCCTCTCCCATTTCAAACTC-3'

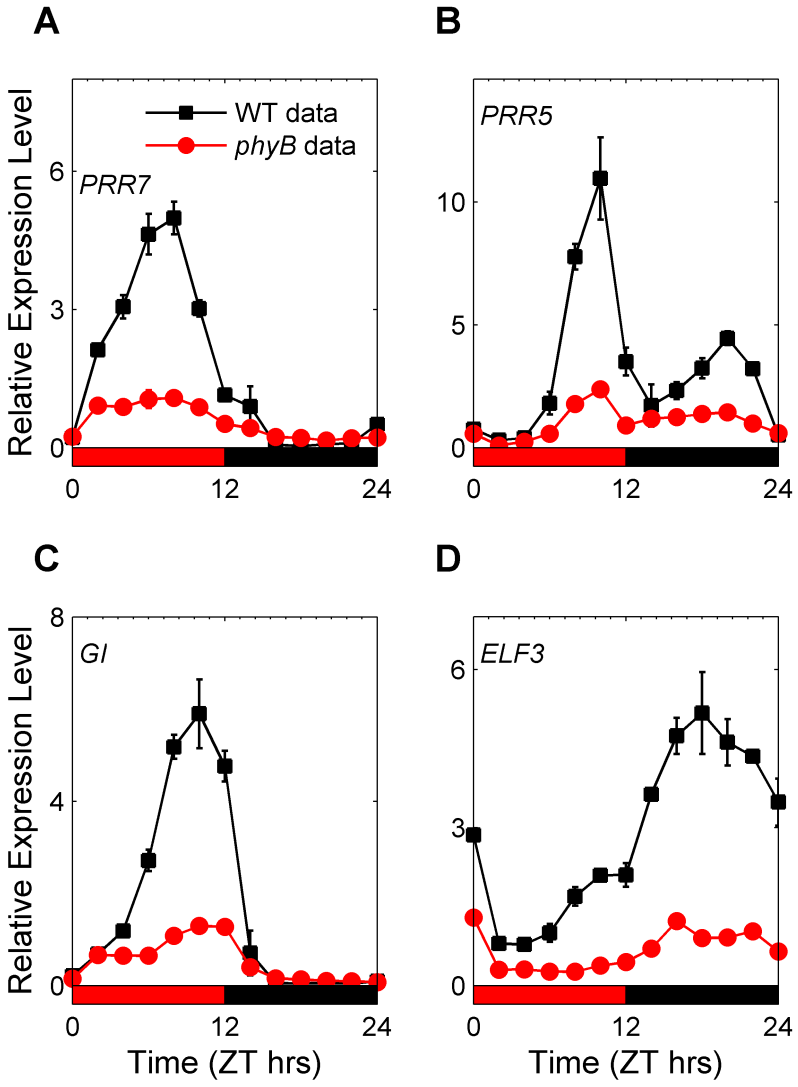


Figure C.1: Effects of *phyB* loss-of-function mutations on clock components at 22°C. Expression rhythms of (A) *PRR7*, (B) *PRR5*, (C) *GI*, and (D) *ELF3* transcription from red light 12L:12D diurnal cycles in WT (black line, filled squares) and *phyB* (red line, filled circles) loss-of-functions mutants. Error bars represent standard error.

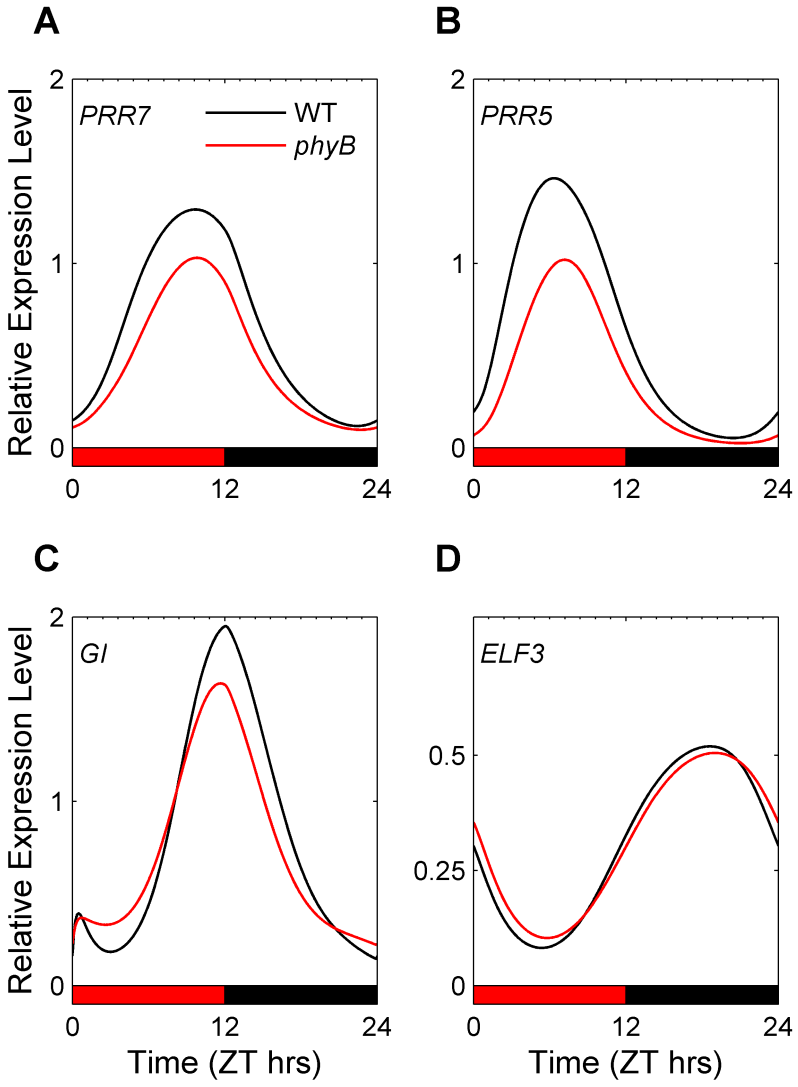


Figure C.2: Simulations of additional clock components in WT and *phyB* at 22°C. As in Figure C.1 with model simulations of WT (black lines) and *phyB* (red lines).

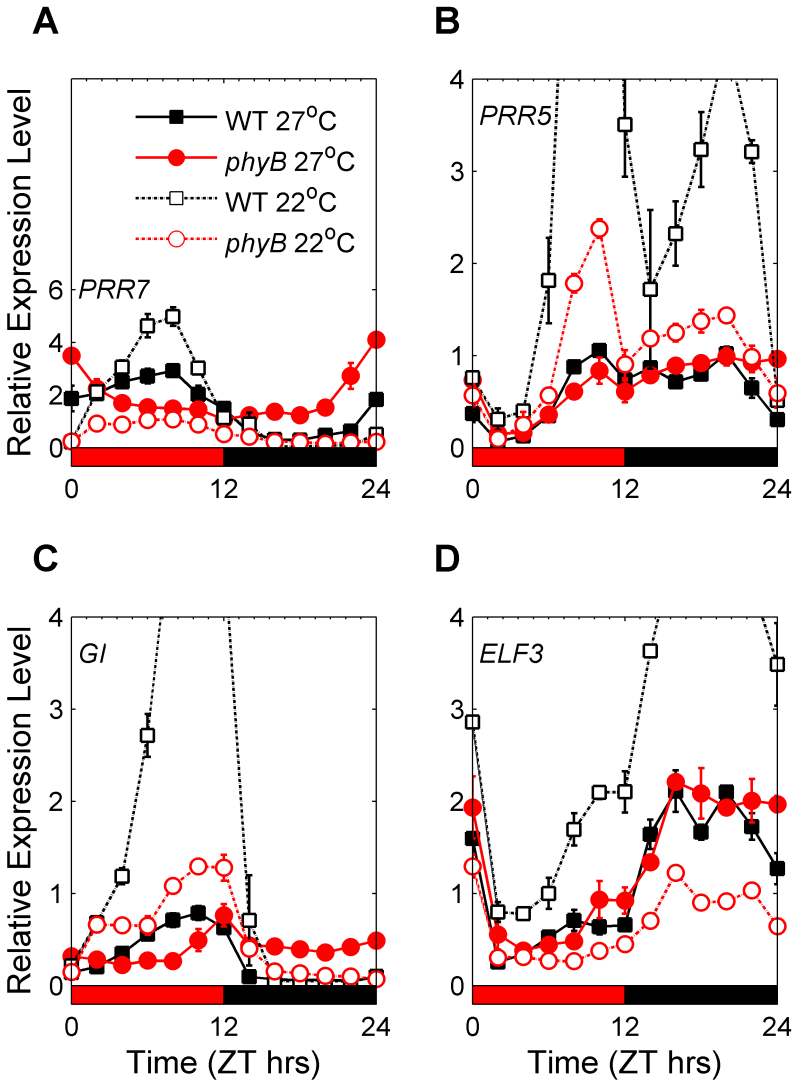


Figure C.3: Effects of *phyB* loss-of-function mutations on clock components at 27°C. As in Figure C.1 for transcriptional rhythms at 27°C.

Bibliography

[1] M. T. Burrows, D. S. Schoeman, L. B. Buckley, P. Moore, E. S. Poloczanska, K. M. Brander, C. Brown, J. F. Bruno, C. M. Duarte, B. S. Halpern, J. Holding, C. V Kappel, W. Kiessling, M. I. O'Connor, J. M. Pandolfi, C. Parmesan, F. B. Schwing, W. J. Sydeman, A. J. Richardson (2011) The pace of shifting climate in marine and terrestrial ecosystems. *Science*, vol. 334, pp. 652–5.

[2] F. Tao, Z. Zhang, W. Shi, Y. Liu, D. Xiao, S. Zhang, Z. Zhu, M. Wang, F. Liu (2013) Single rice growth period was prolonged by cultivars shifts, but yield was damaged by climate change during 1981-2009 in China, and late rice was just opposite. *Glob. Change Biol.*, vol. 19, pp. 3200–9.

[3] H.-Y. Kim, J. Ko, S. Kang, J. Tenhunen (2013) Impacts of climate change on paddy rice yield in a temperate climate. *Glob. Change Biol.*, vol. 19, pp. 548–62.

[4] R. Henriques and P. Mas (2013) Chromatin remodeling and alternative splicing: pre- and post-transcriptional regulation of the Arabidopsis circadian clock. *Semin. Cell Dev. Biol.*, vol. 24, pp. 399–406.

[5] A. Millar, I. Carré, C. Strayer (1995) Circadian clock mutants in Arabidopsis identified by luciferase imaging. *Science*, vol. 267, pp. 1161-3.

[6] P. D. Gould, J. C. W. Locke, C. Larue, M. M. Southern, S. J. Davis, S. Hanano, R. Moyle, R. Milich, J. Putterill, A. J. Millar, A. Hall (2006) The molecular basis of temperature compensation in the Arabidopsis circadian clock. *Plant Cell*, vol. 18, pp. 1177–87.

[7] P. Salomé, D. Weigel, C. R. McClung (2010) The role of the Arabidopsis morning loop components CCA1, LHY, PRR7, and PRR9 in temperature compensation. *Plant Cell*, vol. 22, pp. 3650–61.

[8] P. D. Gould, N. Ugarte, M. Domijan, M. Costa, J. Foreman, D. MacGregor, K. Rose, J. Griffiths, A. J. Millar, B. Finkenstädt, S. Penfield, D. Rand, K. J. Halliday, A. J. W. Hall (2013) Network balance via CRY signalling controls the Arabidopsis circadian clock over ambient temperatures. *Mol. Syst. Biol.*, vol. 9, pp. 1–13.

[9] D. E. Somers, P. F. Devlin, S. A. Kay (1998) Phytochromes and cryptochromes in the entrainment of the Arabidopsis circadian clock. *Science*, vol. 282, pp. 1488–90.

[10] P. F. Devlin and S. A. Kay (2000) Cryptochromes Are Required for Phytochrome Signaling to the Circadian Clock but Not for Rhythmicity. *Plant Cell*, vol. 12, pp. 2499–510.

[11] A. N. Dodd, N. Salathia, A. Hall, E. Kévei, R. Tóth, F. Nagy, J. M. Hibberd, A. J. Millar, A. A. R. Webb (2005) Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science*, vol. 309, pp. 630–3.

[12] A. Pokhilko, A. P. Fernández, K. D. Edwards, M. M. Southern, K. J. Halliday, A. J. Millar (2012) The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops. *Mol. Syst. Biol.*, vol. 8, no. 574.

[13] D. A. Rand, B. V. Shulgin, D. Salazar, A. J. Millar (2004) Design principles underlying circadian clocks. *J. R. Soc. Interface*, vol. 1, pp. 119–30.

[14] O. E. Akman, D. A. Rand, P. E. Brown, A. J. Millar (2010) Robustness from flexibility in the fungal circadian clock. *BMC Syst. Biol.*, vol. 4, no. 88.

[15] Z. Y. Wang and E. M. Tobin (1998) Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell*, vol. 93, pp. 1207–17.

[16] R. Schaffer, N. Ramsay, A. Samach, S. Corden, J. Putterill, I. A. Carre, J. I. Centre, C. Lane (1998) The late elongated hypocotyl mutation of Arabidopsis disrupts circadian rhythms and the photoperiodic control of flowering. *Cell*, vol. 93, pp. 1219–29.

[17] R. M. Green and E. M. Tobin (1999) Loss of the circadian clock-associated protein 1 in Arabidopsis results in altered clock-regulated gene expression. *Proc. Nat. Acad. Sci. USA*, vol. 96, pp. 4176–9.

[18] T. Mizoguchi, K. Wheatley, Y. Hanzawa, L. Wright, M. Mizoguchi, H. Song, I. A. Carre, G. Coupland (2002) LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis. *Cell*, vol. 2, pp. 629–41.

[19] S. L. Harmer and S. A. Kay (2005) Positive and negative factors confer phase-specific circadian regulation of transcription in Arabidopsis. *Plant Cell*, vol. 17, pp. 1926–40.

[20] S. L. Harmer, J. B. Hogenesch, M. Straume, H. S. Chang, B. Han, T. Zhu, X. Wang, J. A. Kreps, S. A. Kay (2000) Orchestrated transcription of key pathways in Arabidopsis by the circadian clock. *Science*, vol. 290, pp. 2110–3.

[21] D. Alabadi, T. Oyama, M. J. Yanovsky, F. G. Harmon, P. Más, S. A. Kay (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. *Science*, vol. 293, no. 5531, pp. 880–3.

[22] M. R. Doyle, S. J. Davis, R. Bastow, H. G. McWatters, L. Kozma-Bognar, A. J. Millar, R. M. Amasino (2002) The ELF4 gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. *Cell*, vol. 419, pp. 9–12.

[23] J. C. W. Locke, M. M. Southern, L. Kozma-Bognár, V. Hibberd, P. E. Brown, M. S. Turner, A. J. Millar (2005) Extension of a genetic network model by iterative experimentation and mathematical analysis. *Mol. Syst. Biol.*, vol. 1, no. 2005.0013.

[24] S. P. Hazen, T. F. Schultz, J. L. Pruneda-Paz, J. O. Borevitz, J. R. Ecker, S. A. Kay (2005) LUX ARRHYTHMO encodes a myb domain protein essential for circadian rhythms. *Proc. Nat. Acad. Sci. USA*, vol. 102, pp. 10387–92.

[25] D. A. Nusinow, A. Helfer, E. E. Hamilton, J. J. King, T. Imaizumi, T. F. Schultz, E. M. Farré, S. A. Kay (2011) The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature*, vol. 475, pp. 398–402.

[26] E. Herrero, E. Kolmos, N. Bujdoso, Y. Yuan, M. Wang, M. C. Berns, H. Uhlworm, G. Coupland, R. Saini, M. Jaskolski, A. Webb, J. Gonçalves, S. J. Davis (2012) EARLY FLOWERING4 recruitment of EARLY FLOWERING3 in the nucleus sustains the Arabidopsis circadian clock. *Plant Cell*, vol. 24, pp. 428–43.

[27] L. Dixon, K. Knox, L. Kozma-Bognar, M. M. Southern, A. Pokhilko, A. J. Millar (2011) Temporal repression of core circadian genes is mediated through EARLY FLOWERING 3 in Arabidopsis. *Curr. Biol.*, vol. 21, pp. 120–5.

[28] A. Helfer, D. A. Nusinow, B. Y. Chow, A. R. Gehrke, M. L. Bulyk, S. A. Kay (2011) LUX ARRHYTHMO encodes a nighttime repressor of circadian gene expression in the Arabidopsis core clock. *Curr. Biol.*, vol. 21, pp. 126–33.

[29] N. Nakamichi, T. Kiba, R. Henriques, T. Mizuno, N.-H. Chua, and H. Sakakibara (2010) PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock. *Plant Cell*, vol. 22, pp. 594–605.

[30] W. Huang, P. Pérez-García, A. Pokhilko, A. J. Millar, I. Antoshechkin, J. L. Riechmann, P. Mas (2012) Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator. *Science*, vol. 75, pp. 75–9.

[31] N. Nakamichi, T. Kiba, M. Kamioka, T. Suzuki, T. Yamashino, T. Higashiyama, H. Sakakibara, T. Mizuno (2012) Transcriptional repressor PRR5 directly regulates clock-output pathways. *Proc. Nat. Acad. Sci. USA*, vol. 109, pp. 17123–8.

[32] P. Mas, W. Kim, D. E. Somers, S. A. Kay (2003) Targeted degradation of TOC1 by ZTL modulates circadian function in *Arabidopsis thaliana*. *Nature* vol. 426, pp. 567–70.

[33] W.-Y. Kim, S. Fujiwara, S.-S. Suh, J. Kim, Y. Kim, L. Han, K. David, J. Putterill, H. G. Nam, D. E. Somers (2007) ZEITLUPE is a circadian photoreceptor stabilized by GIGANTEA in blue light. *Nature*, vol. 449, pp. 356–60.

[34] J.-W. Yu, V. Rubio, N.-Y. Lee, S. Bai, S.-Y. Lee, S.-S. Kim, L. Liu, Y. Zhang, M. L. Irigoyen, J. A. Sullivan, Y. Zhang, I. Lee, Q. Xie, N.-C. Paek, X. W. Deng (2008) COP1 and ELF3 control circadian function and photoperiodic flowering by regulating GI stability. *Mol. Cell*, vol. 32, pp. 617–30.

[35] Y. Kim, J. Lim, M. Yeom, H. Kim, J. Kim, L. Wang, W. Y. Kim, D. E. Somers, H. G. Nam (2013) ELF4 regulates GIGANTEA chromatin access through subnuclear sequestration. *Cell Reports*, vol. 3, pp. 671–7.

[36] J. C. W. Locke, L. Kozma-Bognár, P. D. Gould, B. Fehér, É. Kevei, F. Nagy, M. S. Turner, A. Hall, A. J. Millar (2006) Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. *Mol. Syst. Biol.*, vol. 2, no. 59.

[37] M. N. Zeilinger, E. M. Farré, S. R. Taylor, S. A. Kay, F. J. Doyle (2006) A novel computational model of the circadian clock in Arabidopsis that incorporates PRR7 and PRR9. *Mol. Syst. Biol.*, vol. 2, no. 58.

[38] A. Pokhilko, S. K. Hodge, K. Stratford, K. Knox, K. D. Edwards, A. W. Thomson, T. Mizuno, A. J. Millar (2010) Data assimilation constrains new connections and components in a complex, eukaryotic circadian clock model. *Mol. Syst. Biol.*, vol. 6, no. 416.

[39] A. Pokhilko, P. Mas, A. J. Millar (2013) Modelling the widespread effects of TOC1 signalling on the plant circadian clock and its outputs," *BMC Syst. Biol.*, vol. 7, no. 23.

[40] M. L. Guerriero, A. Pokhilko, A. P. Fernández, K. J. Halliday, A. J. Millar, J. Hillston (2011) Stochastic properties of the plant circadian clock. *J. R. Soc. Interface*, vol. 9, pp. 744–56.

[41] A. Hall, R. M. Bastow, S. J. Davis, S. Hanano, H. G. McWatters, V. Hibberd, M. R. Doyle, S. Sung, K. J. Halliday, R. M. Amasino, A. J. Millar (2003) The TIME FOR COFFEE gene maintains the amplitude and timing of Arabidopsis circadian clocks. *Plant Cell*, vol. 15, pp. 2719–29.

[42] Z. Ding, A. J. Millar, A. M. Davis, S. J. Davis (2007) TIME FOR COFFEE encodes a nuclear regulator in the *Arabidopsis thaliana* circadian clock. *Plant Cell*, vol. 19, pp. 1522–36.

[43] J. Shin, K. Heidrich, A. Sanchez-Villarreal, J. E. Parker, S. J. Davis (2012) TIME FOR COFFEE represses accumulation of the MYC2 transcription factor to provide time-of-day regulation of jasmonate signaling in Arabidopsis. *Plant Cell*, vol. 24, pp. 2470-82.

[44] B. Farinas and P. Mas (2011) Functional implication of the MYB transcription factor RVE8/LCL5 in the circadian control of histone acetylation. *Plant J.*, vol. 66, pp. 318–29.

[45] R. Rawat, N. Takahashi, P. Y. Hsu, M. A. Jones, J. Schwartz, M. R. Salemi, B. S. Phinney, S. L. Harmer (2011) REVEILLE8 and PSEUDO-REPONSE REGULATOR5 form a negative feedback loop within the Arabidopsis circadian clock. *PLoS Genet.*, vol. 7, no. e1001350.

[46] P. Hsu, U. Devisetty, S. Harmer (2013) Accurate timekeeping is controlled by a cycling activator in Arabidopsis. *eLife*, vol. 2, no. e00473.

[47] R. Rawat and J. Schwartz (2009) REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proc. Nat. Acad. Sci. USA*, vol. 106, pp. 16883-8.

[48] J. Pruneda-Paz, G. Breton, A. Para, S. Kay (2009) A functional genomics approach reveals CHE as a component of the Arabidopsis circadian clock. *Science*, vol. 323, pp. 1481–5.

[49] M. L. Rugnone, A. Faigón Soverna, S. E. Sanchez, R. G. Schlaen, C. E. Hernando, D. K. Seymour, E. Mancini, A. Chernomoretz, D. Weigel, P. Más, M. J. Yanovsky (2013) LNK genes integrate light and clock signaling networks at the core of the Arabidopsis oscillator. *Proc. Nat. Acad. Sci. USA*, vol. 110, pp. 12120–5.

[50] N. Nakamichi, M. Kita, K. Niinuma, S. Ito, T. Yamashino, T. Mizoguchi, T. Mizuno (2007) Arabidopsis clock-associated pseudo-response regulators PRR9, PRR7 and PRR5 coordinately and positively regulate flowering time through the canonical CONSTANS-dependent photoperiodic pathway. *Plant Cell Phys*, vol. 48, pp. 822–32.

[51] L. Wang, J. Kim, D. E. Somers (2013) Transcriptional corepressor TOPLESS complexes with pseudoresponse regulator proteins and histone deacetylases to regulate circadian transcription. *Proc. Nat. Acad. Sci. USA*, vol. 110, pp. 761–6.

[52] J. K. Kim and D. B. Forger (2012) A mechanism for robust circadian timekeeping via stoichiometric balance. *Mol. Syst. Biol.*, vol. 8, no. 630.

[53] D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, F. Naef (2011) Mammalian genes are transcribed with widely different bursting kinetics. *Science*, vol. 332, pp. 472–4.

[54] M. Stratmann, D. M. Suter, N. Molina, F. Naef, U. Schibler (2012) Circadian Dbp transcription relies on highly dynamic BMAL1-CLOCK interaction with E boxes and requires the proteasome. *Mol. Cell*, vol. 48, pp. 277–87.

[55] Y. Lee, R. Chen, H. Lee, C. Lee (2011) Stoichiometric relationship among clock proteins determines robustness of circadian rhythms. *J. Biol. Chem.*, vol. 286, pp. 7033–42.

[56] C. Troein, F. Corellou, L. E. Dixon, G. van Ooijen, J. S. O'Neill, F.-Y. Bouget, A. J. Millar (2011) Multiple light inputs to a simple clock circuit allow complex biological rhythms. *Plant J.*, vol. 66, pp. 375–85.

[57] Q. Thommen, B. Pfeuty, F. Corellou, F.-Y. Bouget, M. Lefranc (2012) Robust and flexible response of the Ostreococcus tauri circadian clock to light/dark cycles of varying photoperiod. *FEBS J.*, vol. 279, pp. 3432–48.

[58] M. J. Rust, J. S. Markson, W. S. Lane, D. S. Fisher, E. K. O'Shea (2007) Ordered phosphorylation governs oscillation of a three-protein circadian clock. *Science*, vol. 318, pp. 809–12.

[59] N. Hosokawa, H. Kushige, H. Iwasaki (2013) Attenuation of the posttranslational oscillator via transcription-translation feedback enhances circadian-phase shifts in Synechococcus. *Proc. Nat. Acad. Sci. USA*, vol. 110, pp. 14486-91.

[60] R. S. Edgar, E. W. Green, Y. Zhao, G. Van Ooijen, M. Olmedo, X. Qin, Y. Xu, M. Pan, U. K. Valekunja, K. a Feeney, E. S. Maywood, M. H. Hastings, N. S. Baliga, M. Merrow, A. J. Millar, C. H. Johnson, C. P. Kyriacou, J. S. O'Neill, A. B. Reddy (2012) Peroxiredoxins are conserved markers of circadian rhythms. *Nature*, vol. 485, pp. 459–64.

[61] K. Nozue, M. F. Covington, P. D. Duek, S. Lorrain, C. Fankhauser, S. L. Harmer, J. N. Maloof (2007) Rhythmic growth explained by coincidence between internal and external cues. *Nature*, vol. 448, pp. 358–61.

[62] T. Legnaioli, J. Cuevas, P. Mas (2009) TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. *EMBO J.*, vol. 28, pp. 3745–57.

[63] A. Graf, A. Schlereth, M. Stitt, A. M. Smith (2010) Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proc. Nat. Acad. Sci. USA*, vol. 107, pp. 9458–63.

[64] M. Dong, E. M. Farré, M. F. Thomashow (2011) Circadian clock-associated 1 and late elongated hypocotyl regulate expression of the C-repeat binding factor (CBF) pathway in Arabidopsis. *Proc. Nat. Acad. Sci. USA*, vol. 108, pp. 7241–6.

[65] Y. Nomoto, S. Kubozono, T. Yamashino, N. Nakamichi, T. Mizuno (2012) Circadian clock and PIF4-controlled plant growth: a coincidence mechanism directly integrates a hormone signaling network into the photoperiodic control of plant architectures in Arabidopsis. *Plant Cell Physiol.*, vol. 53, pp. 1950–64.

[66] K. P. Wright, A. W. McHill, B. R. Birks, B. R. Griffin, T. Rusterholz, E. D. Chinoy (2013) Entrainment of the human circadian clock to the natural light-dark cycle. *Curr. Biol.*, vol. 23, pp. 1554-8.

[67] R. Hayama, T. Izawa, K. Shimamoto (2002) Isolation of rice genes possibly involved in the photoperiodic control of flowering by a fluorescent differential display method. *Plant Cell Physiol.*, vol. 43, pp. 494–504.

[68] R. Hayama, G. Coupland, E. Bu (2004) The molecular basis of diversity in the photoperiodic flowering responses of Arabidopsis and rice. *Plant Physiol.*, vol. 135, pp. 677–84.

[69] M. Murakami, Y. Tago, T. Yamashino, T. Mizuno (2007) Comparative overviews of clock-associated genes of *Arabidopsis thaliana* and *Oryza sativa*. *Plant Cell Physiol.*, vol. 48, pp. 110–21.

[70] Y. Nakamura, T. Kato, T. Yamashino, M. Murakami, T. Mizuno (2007) Characterization of a set of Phytochrome-Interacting Factor-Like bHLH proteins in *Oryza sativa*. *Biosci. Biotechnol. Biochem.*, vol. 71, pp. 1183–91.

[71] T. Yamashino, S. Yamawaki, E. Hagui, H. Ueoka-Nakanishi, N. Nakamichi, S. Ito, T. Mizuno (2013) Clock-controlled and FLOWERING LOCUS T (FT)-dependent photoperiodic pathway in *Lotus japonicus* I: Verification of the flowering-associated function of an FT homolog. *Biosci. Biotechnol. Biochem.*, vol. 77, pp. 747-53.

[72] L. Corbesier, I. Gadisseur, G. Silvestre, A. Jacqmard, G. Bernier (1996) Design in *Arabidopsis thaliana* of a synchronous system of floral induction by one long day. *Plant J.*, vol. 9, pp. 947–52.

[73] P. Suárez-López, K. Wheatley, F. Robson, H. Onouchi, F. Valverde, G. Coupland (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. *Nature*, vol. 410, pp. 1116–20.

[74] T. Imaizumi, H. Tran, T. Swartz, W. Briggs, S. Kay (2003) FKF1 is essential for photoperiodic-specific light signalling in Arabidopsis. *Nature*, vol. 426, pp. 302–6.

[75] L. Corbesier, C. Vincent, S. Jang, F. Fornara, Q. Fan, I. Searle, A. Giakountis, S. Farrona, L. Gissot, C. Turnbull, G. Coupland (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science*, vol. 316, pp. 1030–3.

[76] M. Yanovsky, and S. Kay (2002) Molecular basis of seasonal time measurement in Arabidopsis. *Nature*, vol. 419, pp. 308-12.

[77] W. Kim, K. A. Hicks, D. E. Somers (2005) Independent roles for EARLY FLOWERING 3 and ZEITLUPE in the control of circadian timing, hypocotyl length, and flowering time. *Plant Physiol.* vol. 139, pp. 1557–69.

[78] Y. Niwa, S. Ito, N. Nakamichi, T. Mizoguchi, K. Niinuma, T. Yamashino, T. Mizuno (2007) Genetic linkages of the circadian clock-associated genes, TOC1, CCA1 and LHY, in the photoperiodic control of flowering time in *Arabidopsis thaliana*. *Plant Cell Physiol.*, vol. 48, pp. 925–37.

[79] E. Bünning (1936) Die endogene Tagesrhythmik als Grundlage der photoperiodischen reaktion," *Berichte der Deutschen botanischen Gesellschaft*, vol. 54, pp. 590–607.

[80] K. D. Edwards, O. E. Akman, K. Knox, P. J. Lumsden, A. W. Thomson, P. E. Brown, A. Pokhilko, L. Kozma-Bognar, F. Nagy, D. A. Rand, A. J. Millar (2010) Quantitative analysis of regulatory flexibility under changing environmental conditions. *Mol. Syst. Biol.*, vol. 6, no. 424.

[81] M. Sawa, D. A. Nusinow, S. A. Kay, T. Imaizumi (2007) FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science*, vol. 318, pp. 261–5.

[82] J. D. Salazar, T. Saithong, P. E. Brown, J. Foreman, J. C. W. Locke, K. J. Halliday, I. A. Carré, D. A. Rand, A. J. Millar (2009) Prediction of photoperiodic regulators from quantitative gene circuit models. *Cell*, vol. 139, pp. 1170–9.

[83] F. Fornara, K. C. S. Panigrahi, L. Gissot, N. Sauerbrunn, M. Rühl, J. A. Jarillo, G. Coupland (2009) Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. *Dev. Cell*, vol. 17, pp. 75–86.

[84] S. Y. Yoo, Y. Kim, S. Y. Kim, J. S. Lee, J. H. Ahn (2007) Control of flowering time and cold response by a NAC-domain protein in Arabidopsis. *PLoS ONE*, vol. 2, no. e642.

[85] S. Ito, Y. H. Song, A. R. Josephson-Day, R. J. Miller, G. Breton, R. G. Olmstead, T. Imaizumi (2012) FLOWERING BHLH transcriptional activators control expression of the photoperiodic flowering regulator CONSTANS in Arabidopsis. *Proc. Nat. Acad. Sci. USA*, vol. 109, pp. 3582–7.

[86] F. Valverde, A. Mouradov, W. Soppe, D. Ravenscroft, A. Samach, G. Coupland (2004) Photoreceptor regulation of CONSTANS protein in photoperiodic flowering. *Science*, vol. 303, pp. 1003–6.

[87] S. Jang, V. Marchal, K. Panigrahi, S. Wenkel (2008) COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic. *EMBO J.*, vol. 27, pp. 1277–88.

[88] L.-J. Liu, Y.-C. Zhang, Q.-H. Li, Y. Sang, J. Mao, H.-L. Lian, L. Wang, H.-Q. Yang (2008) COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. *Plant Cell*, vol. 20, pp. 292–306.

[89] S. Mangan and U. Alon (2003) Structure and function of the feed-forward loop network motif. *Proc. Nat. Acad. Sci. USA*, vol. 100, pp. 11980-5.

[90] T. Imaizumi, T. F. Schultz, F. G. Harmon, L. A. Ho, S. A. Kay (2005) FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science*, vol. 309, pp. 293–7.

[91] M. Sawa and S. A. Kay (2011) GIGANTEA directly activates Flowering Locus T in *Arabidopsis thaliana*. *Proc. Nat. Acad. Sci. USA*, vol. 108, pp. 11698-703.

[92] L. Rizzini, J.-J. Favory, C. Cloix, D. Faggionato, A. O'Hara, E. Kaiserli, R. Baumeister, E. Schäfer, F. Nagy, G. I. Jenkins, R. Ulm (2011) Perception of UV-B by the Arabidopsis UVR8 protein. *Science*, vol. 332, pp. 103–6.

[93] J. J. Casal (2013) Photoreceptor signaling networks in plant responses to shade. *Annu. Rev. Plant Biol.*, vol. 64, pp. 403-27.

[94] S. Ito, Y. H. Song, T. Imaizumi (2012) LOV domain-containing F-Box proteins: light-dependent protein degradation modules in Arabidopsis. *Mol. Plant*, vol. 5, pp. 573-82.

[95] D. E. Somers, T. F. Schultz, M. Milnamow, S. A. Kay (2000) ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. *Cell*, vol. 101, pp. 319–29.

[96] D. C. Nelson, J. Lasswell, L. E. Rogg, M. A. Cohen, B. Bartel (2000) FKF1, a clock-controlled gene that regulates the transition to flowering in Arabidopsis. *Cell*, vol. 101, pp. 331–40.

[97] J. A. Jarillo, J. Capel, R. H. Tang, H. Q. Yang, J. M. Alonso, J. R. Ecker, and a R. Cashmore (2001) An Arabidopsis circadian clock component interacts with both CRY1 and phyB. *Nature*, vol. 410, pp. 487–90.

[98] T. F. Schultz, T. Kiyosue, M. Yanovsky, M. Wada, S. A. Kay (2001) A Role for LKP2 in the Circadian Clock of Arabidopsis. *Plant Cell*, vol. 13, pp. 2659–70.

[99] Y. Zhu, J. M. Tepperman, C. D. Fairchild, P. H. Quail (2000) Phytochrome B binds with greater apparent affinity than phytochrome A to the basic helix–loop–helix factor PIF3 in a reaction requiring the PAS domain of PIF3. *Proc. Nat. Acad. Sci. USA*, vol. 97, pp. 13419–24.

[100] E. Kevei, P. Gyula, A. Hall, L. Kozma-Bognar, W.-Y. Kim, M. E. Eriksson, R. Toth, S. Hanano, B. Feher, M. M. Southern, R. M. Bastow, A. Viczian, V. Hibberd, S. J. Davis, D. E. Somers, F. Nagy, A. J. Millar (2006) Forward genetic analysis of the circadian clock separates the multiple functions of ZEITLUPE. *Plant Physiol.*, vol. 140, pp. 933–45.

[101] L. Han, M. Mason, E. P. Risseeuw, W. L. Crosby, D. E. Somers (2004) Formation of an SCF(ZTL) complex is required for proper regulation of circadian timing. *Plant J.*, vol. 40, pp. 291–301.

[102] M. Yasuhara, S. Mitsui, H. Hirano, R. Takanabe, Y. Tokioka, N. Ihara, A. Komatsu, M. Seki, K. Shinozaki, T. Kiyosue (2004) Identification of ASK and clock-associated proteins as molecular partners of LKP2 (LOV kelch protein 2) in Arabidopsis. *J. Exp. Bot.*, vol. 55, pp. 2015–27.

[103] The Umiprot Consortium (2013) Update on activities at the Universal Protein Resource (UniProt) in 2013. *Nuc. acids res.*, vol. 41, pp. D43–7.

[104] L. Krall and J. W. Reed (2000) The histidine kinase-related domain participates in phytochrome B function but is dispensable. *Proc. Nat. Acad. Sci. USA*, vol. 97, pp. 8169–74.

[105] Y. Su and J. C. Lagarias (2007) Light-independent phytochrome signaling mediated by dominant GAF domain tyrosine mutants of Arabidopsis phytochromes in transgenic plants. *Plant Cell*, vol. 19, pp. 2124–39.

[106] M. Ni, J. M. Tepperman, P. H. Quail (1998) PIF3, a phytochrome-interacting factor necessary for normal photoinduced signal transduction, is a novel basic helix-loop-helix protein. *Cell*, vol. 95, pp. 657–67.

[107] T. Kiba, R. Henriques, H. Sakakibara, N.-H. Chua (2007) Targeted degradation of PSEUDO-RESPONSE REGULATOR5 by an SCF(ZTL) complex regulates clock function and photomorphogenesis in *Arabidopsis thaliana*. *Plant Cell*, vol. 19, pp. 2516–30.

[108] J. Kim, R. Geng, R. a Gallenstein, D. E. Somers (2013) The F-box protein ZEITLUPE controls stability and nucleocytoplasmic partitioning of GIGANTEA. *Development.*, vol. 140, pp. 4060-9.

[109] A. Baudry, S. Ito, Y. H. Song, A. A. Strait, T. Kiba, S. Lu, R. Henriques, J. L. Pruneda-Paz, N.-H. Chua, E. M. Tobin, S. A. Kay, T. Imaizumi (2010) F-Box proteins FKF1 and LKP2 act in concert with ZEITLUPE to control Arabidopsis clock progression. *Plant Cell*, vol. 22, pp. 606–22.

[110] T. Takase, Y. Nishiyama, H. Tanihigashi, Y. Ogura, Y. Miyazaki, Y. Yamada, T. Kiyosue (2011) LOV KELCH PROTEIN2 and ZEITLUPE repress Arabidopsis photoperiodic flowering under non-inductive conditions, dependent on FLAVIN-BINDING KELCH REPEAT F-BOX1. *Plant J.*, vol. 67, pp. 608–21.

[111] Y. Fukamatsu, S. Mitsui, M. Yasuhara, Y. Tokioka, N. Ihara, S. Fujita, and T. Kiyosue (2005) Identification of LOV KELCH PROTEIN2 (LKP2)-interacting factors that can recruit LKP2 to nuclear bodies. *Plant Cell Physiol.*, vol. 46, pp. 1340–9.

[112] S. Mathews and R. A. Sharrock (1997) Phytochrome gene diversity. *Plant Cell Environ.*, vol. 20, pp. 666–71.

[113] H. Smith (2000) Phytochromes and light signal perception by plants - an emerging synthesis. *Nature*, vol. 407, pp. 585–91.

[114] K. A. Franklin and P. H. Quail (2010) Phytochrome functions in Arabidopsis development. *J. Exp. Bot.*, vol. 61, pp. 11–24.

[115] A. L. Mancinelli (1994) The physiology of phytochrome action. In *Photomorphogenesis in Plants*, R. E. Kendrick and G. M. H. Kronenberg, Dordrecht: Kluwer Academic Publishers, pp. 211–69.

[116] J. Rausenberger, A. Hussong, S. Kircher, D. Kirchenbauer, J. Timmer, F. Nagy, E. Schäfer, C. Fleck (2010) An integrative model for phytochrome B mediated photomorphogenesis: from protein dynamics to physiology. *PLoS ONE*, vol. 5, no. e10721.

[117] S. Kircher, L. Kozma-Bognar, L. Kim, E. Adam, K. Harter, E. Schafer, F. Nagy (1999) Light quality-dependent nuclear import of the plant photoreceptors phytochrome A and B. *Plant Cell*, vol. 11, pp. 1445–56.

[118] R. Yamaguchi, M. Nakamura, N. Mochizuki, S. A. Kay, A. Nagatani (1999) Light-dependent translocation of a phytochrome B-GFP fusion protein to the nucleus in transgenic Arabidopsis. *J. Cell Biol.*, vol. 145, pp. 437–45.

[119] S. Kircher, P. Gil, L. Kozma-bognár, E. Fejes, V. Speth, T. Husselstein-muller, D. Bauer, É. Ádám, E. Schäfer, F. Nagy (2002) Nucleocytoplasmic partitioning of the plant photoreceptors Phytochrome A , B , C , D , and E is regulated differentially by light and exhibits a diurnal rhythm. *Plant Cell*, vol. 14, pp. 1541–55.

[120] T. Kretsch, C. Poppe, E. Scha, È. Freiburg, È. B. Botanik (2000) A new type of mutation in the plant photoreceptor phytochrome B causes loss of photoreversibility and an extremely enhanced light sensitivity. *Plant J.*, vol. 22, pp. 177–86.

[121] M. de Lucas, J. Davière, M. Rodríguez-Falcon, M. Pontin, J. Iglesias-Pedraz, S. Lorrain, C. Fankhauser, M. Blazquez, E. Titarenko, S. Prat (2008) A molecular framework for light and gibberellin control of cell elongation. *Nature*, vol. 451, pp. 480–4.

[122] M. Koini, L. Alvey, T. Allen, C. a Tilley, N. P. Harberd, G. C. Whitelam, K. A. Franklin (2009) High temperature-mediated adaptations in plant architecture require the bHLH transcription factor PIF4. *Curr. Biol.*, vol. 19, pp. 408–13.

[123] J. A. Stavang, J. Gallego-Bartolomé, M. D. Gómez, S. Yoshida, T. Asami, J. E. Olsen, J. L. García-Martínez, D. Alabadí, M. A. Blázquez (2009) Hormonal regulation of temperature-induced growth in Arabidopsis. *Plant J.*, vol. 60, pp. 589–601.

[124] K. A. Franklin, S. H. Lee, D. Patel, S. V. Kumar, A. K. Spartz, C. Gu, S. Ye, P. Yu, G. Breen, J. D. Cohen, P. a Wigge, W. M. Gray (2011) PHYTOCHROME-INTERACTING FACTOR 4 (PIF4) regulates auxin biosynthesis at high temperature. *Proc. Nat. Acad. Sci. USA*, vol. 108, pp. 20231–5.

[125] J. Foreman, H. Johansson, P. Hornitschek, E.-M. Josse, C. Fankhauser, K. J. Halliday (2011) Light receptor action is critical for maintaining plant biomass at warm ambient temperatures. *Plant J.*, vol. 65, pp. 441–52.

[126] K. Nozue, S. L. Harmer, J. N. Maloof (2011) Genomic analysis of circadian clock-, light-, and growth-correlated genes reveals PHYTOCHROME-INTERACTING FACTOR5 as a modulator of auxin signaling in Arabidopsis. *Plant Physiol.*, vol. 156, pp. 357–72.

[127] J. L. Stewart, J. N. Maloof, J. L. Nemhauser (2011) PIF genes mediate the effect of sucrose on seedling growth dynamics. *PLoS ONE*, vol. 6, no. e19894.

[128] C.-M. Lee and M. F. Thomashow (2012) Photoperiodic regulation of the C-repeat binding factor (CBF) cold acclimation pathway and freezing tolerance in Arabidopsis thaliana. *Proc. Nat. Acad. Sci. USA*, vol. 109, pp. 15054–9.

[129] E. Oh, J.-Y. Zhu, Z.-Y. Wang (2012) Interaction between BZR1 and PIF4 integrates brassinosteroid and environmental responses. *Nat. Cell Biol.*, vol. 14, pp. 802–9.

[130] D. Chen, G. Xu, W. Tang, Y. Jing, Q. Ji, Z. Fei, R. Lin (2013) Antagonistic Basic Helix-Loop-Helix/bZIP transcription factors form transcriptional modules that integrate light and reactive oxygen species signaling in Arabidopsis. *Plant Cell*, vol. 25, pp. 1657-73.

[131] P. Hornitschek, S. Lorrain, V. Zoete, O. Michielin, C. Fankhauser (2009) Inhibition of the shade avoidance response by formation of non-DNA binding bHLH heterodimers. *EMBO J.*, vol. 28, pp. 3893–902.

[132] P. Leivar, J. M. Tepperman, M. M. Cohn, E. Monte, B. Al-Sady, E. Erickson, P. H. Quail (2012) Dynamic antagonism between phytochromes and PIF family basic helix-loop-helix factors induces selective reciprocal responses to light and shade in a rapidly responsive transcriptional network in Arabidopsis. *Plant Cell*, vol. 24, pp. 1398–419.

[133] P. Hornitschek, M. V Kohnen, S. Lorrain, J. Rougemont, K. Ljung, I. López-Vidriero, J. M. Franco-Zorrilla, R. Solano, M. Trevisan, S. Pradervand, I. Xenarios, C. Fankhauser (2012) Phytochrome interacting factors 4 and 5 control seedling growth in changing light conditions by directly controlling auxin signaling. *Plant J.*, vol. 71, pp. 699–711.

[134] M. Chen, Y. Tao, J. Lim, A. Shaw, J. Chory (2005) Regulation of phytochrome B nuclear localization through light-dependent unmasking of nuclear-localization signals. *Curr. Biol.*, vol. 15, pp. 637–42.

[135] M. Ni, J. M. Tepperman, P. H. Quail (1999) Binding of phytochrome B to its nuclear signalling partner PIF3 is reversibly induced by light. *Nature*, vol. 400, pp. 781–4.

[136] E. A. Kikis, Y. Oka, M. E. Hudson, A. Nagatani, P. H. Quail (2009) Residues clustered in the light-sensing knot of phytochrome B are necessary for conformer-specific binding to signaling partner PIF3. *PLoS Genet.*, vol. 5, no. e1000352.

[137] B. Al-Sady, E. A. Kikis, E. Monte, P. H. Quail (2008) Mechanistic duality of transcription factor function in phytochrome signaling. *Proc. Nat. Acad. Sci. USA*, vol. 105, pp. 2232–7.

[138] W. Ni, S.-L. Xu, R. J. Chalkley, T. N. D. Pham, S. Guan, D. A. Maltby, A. L. Burlingame, Z.-Y. Wang, P. H. Quail (2013) Multisite light-induced phosphorylation of the transcription factor PIF3 is necessary for both its rapid degradation and concomitant negative feedback modulation of photoreceptor phyB levels in Arabidopsis. *Plant Cell*, vol. 25, pp. 2679-98.

[139] I.-C. Jang, R. Henriques, H. S. Seo, A. Nagatani, N.-H. Chua (2010) Arabidopsis PHYTOCHROME INTERACTING FACTOR proteins promote phytochrome B polyubiquitination by COP1 E3 ligase in the nucleus. *Plant Cell*, vol. 22, pp. 2370–83.

[140] R. A. Sharrock and T. Clack (2002) Patterns of expression and normalized levels of the five Arabidopsis phytochromes. *Plant Physiol.*, vol. 130, pp. 442–56.

[141] E. Park, J. Park, J. Kim, A. Nagatani, J. C. Lagarias, G. Choi (2012) Phytochrome B inhibits binding of phytochrome-interacting factors to their target promoters. *Plant J.*, vol. 72, pp. 537–46.

[142] J. M. Tepperman, M. E. Hudson, R. Khanna, T. Zhu, S. H. Chang, X. Wang, P. H. Quail (2004) Expression profiling of phyB mutant demonstrates substantial contribution of other phytochromes to red-light-regulated gene expression during seedling de-etiolation. *Plant J.*, vol. 38, pp. 725–39.

[143] S. L. Anderson, D. E. Somers, A. J. Millar, K. Hanson, J. Chory, S. A. Kay (1997) Attenuation of phytochrome A and B signaling pathways by the Arabidopsis circadian clock. *Plant Cell*, vol. 9, pp. 1727–43.

[144] S. Ito, A. Matsushika, H. Yamada, S. Sato, T. Kato, S. Tabata, T. Yamashino, T. Mizuno (2003) Characterization of the APRR9 pseudo-response regulator belonging to the APRR1/TOC1 quintet in *Arabidopsis thaliana*. *Plant Cell Physiol.*, vol. 44, pp. 1237–45.

[145] I.-C. Jang, P. J. Chung, H. Hemmes, C. Jung, N.-H. Chua (2011) Rapid and reversible light-mediated chromatin modifications of Arabidopsis phytochrome A locus. *Plant Cell*, vol. 23, pp. 459–70.

[146] W. Hu, K. A. Franklin, R. A. Sharrock, M. A. Jones, S. L. Harmer, J. C. Lagarias (2013) Unanticipated regulatory roles for Arabidopsis phytochromes revealed by null mutant analysis. *Proc. Nat. Acad. Sci. USA*, vol. 110, pp. 1542–7.

[147] H. Guo, H. Yang, T. C. Mockler, C. Lin (1998) Regulation of flowering time by Arabidopsis photoreceptors. *Science*, vol. 279, pp. 1360–3.

[148] T. C. Mockler, H. Guo, H. Yang, H. Duong, C. Lin (1999) Antagonistic actions of Arabidopsis cryptochromes and phytochrome B in the regulation of floral induction. *Development*, vol. 126, pp. 2073–82, 1999.

[149] K. J. Halliday, M. G. Salter, E. Thingnaes, G. C. Whitelam (2003) Phytochrome control of flowering is temperature sensitive and correlates with expression of the floral integrator FT. *Plant J.*, vol. 33, pp. 875–85.

[150] K. Miyata, M. Calviño, A. O. Oda, H. Sugiyama, T. Mizoguchi (2011) Suppression of late-flowering and semi-dwarf phenotypes in the Arabidopsis clock mutant lhy-12;cca1-101 by phyB under continuous light. *Plant Signal. Behav.*, vol. 6, pp. 1162–71.

[151] S. Y. Kim, X. Yu, S. D. Michaels (2008) Regulation of CONSTANS and FLOWERING LOCUS T expression in response to changing light quality. *Plant Physiol.*, vol. 148, pp. 269-79.

[152] A. C. Wollenberg, B. Strasser, P. D. Cerdán, R. M. Amasino (2008) Acceleration of flowering during shade avoidance in Arabidopsis alters the balance between FLOWERING LOCUS C-mediated repression and photoperiodic induction of flowering. *Plant Physiol.*, vol. 148, pp. 1681–94.

[153] Y. H. Song, R. W. Smith, B. J. To, A. J. Millar, T. Imaizumi (2012) FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. *Science*, vol. 336, pp. 1045–9.

[154] S. Penfield (2008) Temperature perception and signal transduction in plants. *New Phytol.*, vol. 179, pp. 615–28.

[155] E. J. Stockinger, S. J. Gilmour, M. F. Thomashow (1997) *Arabidopsis thaliana* CBF1 encodes an AP2 domain-containing transcriptional activator that binds to the C-repeat/DRE, a cis-acting DNA regulatory element that stimulates transcription in response to low temperature and water deficit. *Proc. Nat. Acad. Sci. USA*, vol. 94, pp. 1035–40.

[156] S. G. Fowler, D. Cook, M. F. Thomashow (2005) Low temperature induction of Arabidopsis CBF1, 2, and 3 is gated by the circadian clock. *Plant Physiol.*, vol. 137, pp. 961–8.

[157] H. Lee, L. Xiong, Z. Gong, M. Ishitani, B. Stevenson, J. K. Zhu (2001) The Arabidopsis HOS1 gene negatively regulates cold signal transduction and encodes a RING finger protein that displays cold-regulated nucleo-cytoplasmic partitioning. *Genes Dev.*, vol. 15, pp. 912–24.

[158] C.-H. Dong, M. Agarwal, Y. Zhang, Q. Xie, J.-K. Zhu (2006) The negative regulator of plant cold responses, HOS1, is a RING E3 ligase that mediates the ubiquitination and degradation of ICE1. *Proc. Nat. Acad. Sci. USA*, vol. 103, pp. 8281–6.

[159] K. Miura, J. B. Jin, J. Lee, C. Y. Yoo, V. Stirm, T. Miura, E. N. Ashworth, R. a Bressan, D.-J. Yun, P. M. Hasegawa (2007) SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in Arabidopsis. *Plant Cell*, vol. 19, pp. 1403–14.

[160] N. Nakamichi, M. Kusano, A. Fukushima, M. Kita, S. Ito, T. Yamashino, K. Saito, H. Sakakibara, T. Mizuno (2009) Transcript profiling of an Arabidopsis PSEUDO RESPONSE REGULATOR arrhythmic triple mutant reveals a role for the circadian clock in cold stress response. *Plant Cell Physiol.*, vol. 50, pp. 447–62.

[161] Y. Nomoto, S. Kubozono, M. Miyachi, T. Yamashino, N. Nakamichi, T. Mizuno (2012) A circadian clock and PIF4-mediated double coincidence mechanism is implicated in the thermo-sensitive photoperiodic control of plant architectures in *Arabidopsis thaliana*. *Plant Cell Physiol.*, vol. 53, pp. 1965–73.

[162] T. Yamashino, Y. Nomoto, S. Lorrain, M. Miyachi, S. Ito, N. Nakamichi, C. Fankhauser, T. Mizuno (2013) Verification at the protein level of the PIF4-mediated external coincidence model for the temperature-adaptive photoperiodic control of plant growth in *Arabidopsis thaliana*. *Plant Signal. Behav.*, vol. 8, no. e23390.

[163] J. Keily, D. R. MacGregor, R. W. Smith, A. J. Millar, K. J. Halliday, S. Penfield (2013) Model selection reveals control of cold signalling by evening-phased components of the plant circadian clock. *Plant J.*, vol. 76, 247-57.

[164] E. Karayekov, R. Sellaro, M. Legris, M. J. Yanovsky, J. J. Casal (2013) Heat shock-induced fluctuations in clock and light signaling enhance phytochrome B-mediated Arabidopsis deetiolation. *Plant Cell*, vol. 25, pp. 2892-906.

[165] A. B. James, N. H. Syed, S. Bordage, J. Marshall, G. A. Nimmo, G. I. Jenkins, P. Herzyk, J. W. S. Brown, H. G. Nimmo (2012) Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. *Plant Cell*, vol. 24, pp. 961–81.

[166] X. Wang, F. Wu, Q. Xie, H. Wang, Y. Wang, Y. Yue, O. Gahura, S. Ma, L. Liu, Y. Cao, Y. Jiao, F. Puta, C. R. McClung, X. Xu, L. Ma (2012) SKIP is a component of the spliceosome linking alternative splicing and the circadian clock in Arabidopsis. *Plant Cell*, vol. 24, pp. 3278–95.

[167] P. J. Seo, M.-J. Park, M.-H. Lim, S.-G. Kim, M. Lee, I. T. Baldwin, C.-M. Park (2012) A self-regulatory circuit of CIRCADIAN CLOCK-ASSOCIATED1 underlies the circadian clock regulation of temperature responses in Arabidopsis. *Plant Cell*, vol. 24, pp. 2427–42.

[168] J. Majercak, W. Chen, I. Edery (2004) Splicing of the period gene 3'-terminal intron is regulated by light , circadian clock factors , and phospholipase C. *Mol. Cell. Biol.*, vol. 24, pp. 3359–72.

[169] A. C. R. Diernfellner, T. Schafmeier, M. W. Merrow, M. Brunner (2005) Molecular mechanism of temperature sensing by the circadian clock of *Neurospora crassa*. *Genes Dev.*, vol. 19, pp. 1968–73.

[170] J.-H. Jung, P. J. Seo, C.-M. Park (2012) The E3 ubiquitin ligase HOS1 regulates Arabidopsis flowering by mediating CONSTANS degradation under cold stress. *J. Biol. Chem.*, vol. 287, pp. 43277–87.

[171] J. Lee, J. Kim, S. Kim, H. J. Cho, J. Kim, J. H. Ahn (2012) The E3 ubiquitin ligase HOS1 regulates low ambient temperature-responsive flowering in *Arabidopsis thaliana*. *Plant Cell Physiol.*, vol. 53, pp. 1802–14.

[172] A. Lazaro, F. Valverde, M. Pineiro, J. A. Jarillo (2012) The Arabidopsis E3 ubiquitin ligase HOS1 negatively regulates CONSTANS abundance in the photoperiodic control of flowering. *Plant Cell*, vol. 24, pp. 982–99.

[173] S. Balasubramanian, S. Sureshkumar, J. Lempe, D. Weigel (2006) Potent induction of *Arabidopsis thaliana* flowering by elevated growth temperature. *PLoS Genet.*, vol. 2, no. e106.

[174] S. V. Kumar, D. Lucyshyn, K. E. Jaeger, E. Alós, E. Alvey, N. P. Harberd, P. A. Wigge (2012) Transcription factor PIF4 controls the thermosensory activation of flowering. *Nature*, vol. 484, pp. 242–5.

[175] Y. H. Chew, A. M. Wilczek, M. Williams, S. M. Welch, J. Schmitt, K. J. Halliday (2012) An augmented Arabidopsis phenology model reveals seasonal temperature control of flowering time. *New Phytol.*, vol. 194, pp. 654–65.

[176] K. A. Franklin and G. C. Whitelam (2007) Light-quality regulation of freezing tolerance in *Arabidopsis thaliana*. *Nat. Genet.*, vol. 39, pp. 1410–3.

[177] R. Catala, J. Medina, J. Salinas (2011) Integration of low temperature and light signaling during cold acclimation response in Arabidopsis. *Proc. Nat. Acad. Sci. USA*, vol. 108, pp. 16475–80.

[178] Y. Niwa, T. Yamashino, T. Mizuno (2009) The circadian clock regulates the photoperiodic response of hypocotyl elongation through a coincidence mechanism in *Arabidopsis thaliana*. *Plant Cell Physiol.*, vol. 50, pp. 838–54.

[179] S. X. Lu, C. J. Webb, S. M. Knowles, S. H. J. Kim, Z.-Y. Wang, E. M. Tobin (2011) CCA1 and ELF3 interact in the control of hypocotyl length and flowering time in Arabidopsis. *Plant Physiol.*, vol. 158, pp. 1079–88.

[180] H. Akaike (1974) A new look at the statistical model identification. *IEEE T. Auto. Contr.*, vol. 19, pp. 716–23.

[181] C. M. Hurvich and C.-L. Tsai (1989) Regression and time series model selection in small samples. *Biometrika*, vol. 76, pp. 297–307.

[182] K. Burnham and D. Anderson (2004) Multimodel inference understanding AIC and BIC in model selection. *Sociol. Methods Res.*, vol. 33, pp. 261–304.

[183] N. Friel and A. N. Pettitt (2008) Marginal likelihood estimation via power posteriors. *J. Roy. Statist. Soc. Ser. B*, vol. 70, pp. 589–607.

[184] V. Vyshemirsky and M. A. Girolami (2008) Bayesian ranking of biochemical system models. *Bioinformatics*, vol. 24, pp. 833–9.

[185] K. Vandepoele, M. Quimbaya, T. Casneuf, L. De Veylder, Y. Van De Peer (2009) Unraveling transcriptional control in Arabidopsis using cis-regulatory elements and coexpression networks. *Plant Physiol.*, vol. 150, pp. 535–46.

[186] K. Higo, Y. Ugawa, M. Iwamoto, T. Korenaga (1999) Plant cis-acting regulatory DNA elements (PLACE) database. *Nuc. Acids Res.*, vol. 27, pp. 297–300.

[187] S. K. Palaniswamy, S. James, H. Sun, R. S. Lamb, R. V Davuluri, E. Grotewold (2006) AGRIS and AtRegNet: a platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol.*, vol. 140, pp. 818–29.

[188] M. D. Mikkelsen and M. F. Thomashow (2009) A role for circadian evening elements in cold-regulated gene expression in Arabidopsis. *Plant J.*, vol. 60, pp. 328–39.

[189] T. P. Michael, G. Breton, S. P. Hazen, H. Priest, T. C. Mockler, S. A. Kay, J. Chory (2008) A morning-specific phytohormone gene expression program underlying rhythmic plant growth. *PLoS Biol.*, vol. 6, no. e225.

[190] C. L. Walcher and J. L. Nemhauser (2012) Bipartite promoter element required for auxin response. *Plant Physiol.*, vol. 158, pp. 273–82.

[191] T. C. Mockler, T. P. Michael, H. D. Priest, R. Shen, C. M. Sullivan, S. A. Givan, C. McEntee, S. A. Kay, J. Chory (2007) The DIURNAL project: Diurnal and circadian expression profiling, model-based pattern matching and promoter analysis. *Cold Spring Harb. Symp. Quant. Biol.*, vol. 72, pp. 353–63.

[192] S. M. Hong, S. C. Bahn, A. Lyu, H. S. Jung, J. H. Ahn (2010) Identification and testing of superior reference genes for a starting pool of transcript normalization in Arabidopsis. *Plant Cell Physiol.*, vol. 51, pp. 1694–706.

[193] J. D. Nelson, O. Denisenko, P. Sova, K. Bomsztyk (2006) Fast chromatin immunoprecipitation assay. *Nuc. Acids Res.*, vol. 34, no. e2.

[194] E. Yakir, D. Hilman, I. Kron, M. Hassidim, N. Melamed-Book, R.M. Green (2009) Posttranslational regulation of CIRCADIAN CLOCK ASSOCIATED 1 in the circadian oscillator of Arabidopsis. *Plant Physiol.*, vol. 150, pp. 844-57.

[195] O. S. Lau, X. Huang, J.-B. Charron, J.-H. Lee, G. Li, X. W. Deng (2011) Interaction of Arabidopsis DET1 with CCA1 and LHY in mediating transcriptional repression in the plant circadian clock. *Mol. Cell*, vol. 43, pp. 703–12.

[196] P. Leivar, J. M. Tepperman, E. Monte, R. H. Calderon, T. L. Liu, P. H. Quail (2009) Definition of early transcriptional circuitry involved in light-induced reversal of PIF-imposed repression of photomorphogenesis in young Arabidopsis seedlings. *Plant Cell*, vol. 21, pp. 3535–53.

[197] A. C. Hindmarsh, P. N. Brown, K. E. Grant, S. L. Lee, R. Serban, D. A. N. E. Shumaker, C. S. Woodward (2005) SUNDIALS: suite of nonlinear and differential/algebraic equation solvers. *ACM T. Math. Soft.*, vol. 31, pp. 363–96.

[198] R. Serban and A. C. Hindmarsh (2003) CVODES: an ODE solver with sensitivity analysis capabilities. *ACM T. Math. Soft.*

[199] H. Mühlenbein, M. Schomisch, J. Born (1991) The parallel genetic algorithm as function optimizer. *Parallel Comput.*, vol. 17, pp. 619–32.

[200] K. Brown, C. Hill, G. Calero (2004) The statistical mechanics of complex signaling networks: nerve growth factor signaling. *Phys. Biol.*, vol. 1, pp. 184–95.

[201] K. Laskey and J. Myers (2003) Population markov chain monte carlo. *Mach. Learn.*, vol. 50, pp. 175–96.

[202] J. Benesty, J. Chen, Y. Huang, I. Cohen (2009) Pearson correlation coefficient. In *Noise Reduction in Speech Processing*, Springer, pp. 37–40.

[203] B. Daniels and Y. Chen (2008) Sloppiness, robustness, and evolvability in systems biology. *Curr. Opin. Biotechnol.*, vol. 19, pp. 389–95.

[204] J. Deely and D. Lindley (1981) Bayes empirical Bayes. *J. Amer. Statist. Soc.*, vol. 76, pp. 833–41.

[205] R. E. Kass, and D. Steffey (1989) Approximate Bayesian inference in conditionally independent hierarchical models (parametric empirical Bayes models). *J. Amer. Statist. Soc.*, vol. 84, pp. 717–26.

[206] A. Gelman (2008) Objections to Bayesian statistics. *Bayesian Anal.*, vol. 3, pp. 443–50.

[207] C. S. Pittendrigh (1960) Circadian rhythms and the circadian organization of living systems. *Cold Spring Harb. Symp. Quant. Biol.*, vol. 25, pp. 159–84.

[208] M. Ashyraliyev (2009) Systems biology: parameter estimation for biochemical models. *FEBS J.*, vol. 276, no. 4, pp. 886–902.

[209] S. Ito, Y. Niwa, N. Nakamichi, H. Kawamura, T. Yamashino, T. Mizuno (2008) Insight into missing genetic links between two evening-expressed pseudo-response regulator genes TOC1 and PRR5 in the circadian clock-controlled circuitry in *Arabidopsis thaliana*. *Plant Cell Physiol.*, vol. 49, pp. 201–13.

[210] E. Salvo-Chirnside, S. Kane, L. E. Kerr (2011) Protocol: high throughput silica-based purification of RNA from Arabidopsis seedlings in a 96-well format. *Plant Methods*, vol. 7, no. 40.

[211] A. Pokhilko, J. A. Ramos, H. Holtan, D. R. Maszle, R. Khanna, and A. J. Millar (2011) Ubiquitin ligase switch in plant photomorphogenesis: A hypothesis. *J. Theor. Biol.*, vol. 270, pp. 31–41.

[212] D. E. Somers, W. Kim, R. Geng (2004) The F-box protein ZEITLUPE confers dosage-dependent control on the circadian clock, photomorphogenesis, and flowering time. *Plant Cell*, vol. 16, pp. 769–82.

[213] R. Yoshida, R. Fekih, S. Fujiwara, A. Oda, K. Miyata, Y. Tomozoe, M. Nakagawa, K. Niinuma, K. Hayashi, H. Ezura, G. Coupland, T. Mizoguchi (2009) Possible role of early flowering 3 (ELF3) in clock-dependent floral regulation by short vegetative phase (SVP) in *Arabidopsis thaliana*. *New Phytol.*, vol. 182, pp. 838–50.

[214] J. W. Reed, P. Nagpal, R. M. Bastow, K. S. Solomon, M. J. Dowson-Day, R. P. Elumalai, A. J. Millar (2000) Independent action of ELF3 and phyB to control hypocotyl elongation and flowering time. *Plant Physiol.*, vol. 122, pp. 1149–60.

[215] X. L. Liu, M. F. Covington, C. Fankhauser, J. Chory, D. R. Wagner (2001) ELF3 encodes a circadian clock-regulated nuclear protein that functions in an Arabidopsis PHYB signal transduction pathway. *Plant Cell*, vol. 13, pp. 1293–304.

[216] S. Makino, A. Matsushika, M. Kojima, T. Yamashino, T. Mizuno (2002) The APRR1/TOC1 quintet implicated in circadian rhythms of *Arabidopsis thaliana*: I . Characterization with APRR1-overexpressing plants. *Plant Cell Physiol.*, vol. 43, pp. 58–69.

[217] T. Yamashino, A. Matsushika, T. Fujimori, S. Sato, T. Kato, S. Tabata, T. Mizuno (2003) A link between circadian-controlled bHLH factors and the APRR1/TOC1 quintet in *Arabidopsis thaliana*.," *Plant Cell Physiol.*, vol. 44, pp. 619–29.

[218] C. Andronis, S. Barak, S. M. Knowles, S. Sugano, E. M. Tobin (2008) The clock protein CCA1 and the bZIP transcription factor HY5 physically interact to regulate gene expression in Arabidopsis. *Mol. Plant*, vol. 1, pp. 58–67.

[219] V. Yadav, S. Kundu, D. Chattopadhyay, P. Negi, N. Wei, X.-W. Deng, S. Chattopadhyay (2002) Light regulated modulation of Z-box containing promoters by photoreceptors and downstream regulatory components, COP1 and HY5, in Arabidopsis. *Plant J.*, vol. 31, pp. 741–53.

[220] A. Matsushika, M. Kawamura, Y. Nakamura, T. Kato, M. Murakami, T. Yamashino, T. Mizuno (2007) Characterization of circadian-associated pseudo-response regulators: II. The function of PRR5 and its molecular dissection in *Arabidopsis thaliana*. *Biosci. Biotech. Biochem.*, vol. 71, pp. 535–544.

[221] M. Takase, T. Mizoguchi, T. Kozuka, H. Tsukaya (2013) The unique function of the Arabidopsis circadian clock gene PRR5 in the regulation of shade avoidance response. *Plant Signal. Behav.*, vol. 8, no. e23534.

[222] S. Fujiwara, L. Wang, L. Han, S.-S. Suh, P. A. Salomé, C. R. McClung, D. E. Somers (2008) Post-translational regulation of the Arabidopsis circadian clock through selective proteolysis and phosphorylation of pseudo-response regulator proteins. *J. Biol. Chem.*, vol. 283, pp. 23073–83.

[223] E. Huq, J. M. Tepperman, P. H. Quail (2000) GIGANTEA is a nuclear protein involved in phytochrome signaling in Arabidopsis. *Proc. Nat. Acad. Sci. USA*, vol. 97, pp. 9789–94.

[224] R. Khanna, E. A. Kikis, P. H. Quail (2003) EARLY FLOWERING 4 functions in Phytochrome B-regulated seedling de-etiolation. *Plant Physiol.*, vol. 133, pp. 1530–8.

[225] K. D. Edwards, J. R. Lynn, P. Gyula, F. Nagy, A. J. Millar (2005) Natural allelic variation in the temperature-compensation mechanisms of the *Arabidopsis thaliana* circadian clock. *Genetics*, vol. 170, pp. 387–400.

[226] D. A. Rand (2008) Mapping global sensitivity of cellular network dynamics: sensitivity heat maps and a global summation law. *J. R. Soc. Interface*, vol. 5 Suppl 1, pp. S59–S69.

[227] J. D. Plautz, M. Straume, R. Stanewsky, C. F. Jamison, C. Brandes, H. B. Dowse, J. C. Hall, and S. A. Kay (1997) Quantitative analysis of Drosophila period gene transcription in living animals. *J. Biol. Rhythms*, vol. 12, pp. 204–17.

[228] M. T. Osterlund and X. W. Deng (1998) Multiple photoreceptors mediate the light-induced reduction of GUS-COP1 from Arabidopsis hypocotyl nuclei. *Plant J.*, vol. 16, pp. 201–8.

[229] D. Zhu, A. Maier, J.-H. Lee, S. Laubinger, Y. Saijo, H. Wang, L.-J. Qu, U. Hoecker, X. W. Deng (2008) Biochemical characterization of Arabidopsis complexes containing CONSTITUTIVELY PHOTOMORPHOGENIC1 and SUPPRESSOR OF PHYA proteins in light control of plant development. *Plant Cell*, vol. 20, pp. 2307–23.

[230] H. Chen, X. Huang, G. Gusmaroli, W. Terzaghi, O. S. Lau, Y. Yanagawa, Y. Zhang, J. Li, J.-H. Lee, D. Zhu, X. W. Deng (2010) Arabidopsis CULLIN4-

Damaged DNA Binding Protein 1 interacts with CONSTITUTIVELY PHOTOMORPHOGENIC1-SUPPRESSOR OF PHYA complexes to regulate photomorphogenesis and flowering time. *Plant Cell*, vol. 22, pp. 108–23.

[231] H. An, C. Roussot, P. Suárez-López, L. Corbesier, C. Vincent, M. Piñeiro, S. Hepworth, A. Mouradov, S. Justin, C. Turnbull, G. Coupland (2004) CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of Arabidopsis. *Development*, vol. 131, pp. 3615–26.

[232] S. Wenkel, F. Turck, K. Singer, L. Gissot, J. Le Gourrierec, A. Samach, G. Coupland (2006) CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of Arabidopsis. *Plant Cell*, vol. 18, pp. 2971–84.

[233] J. M. Gendron, J. L. Pruneda-Paz, C. J. Doherty, A. M. Gross, S. E. Kang, S. A. Kay (2012) Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. *Proc. Nat. Acad. Sci. USA*, vol. 109, pp. 3167–72.

[234] K. E. Jaeger, N. Pullen, S. Lamzin, R. J. Morris, and P. A. Wigge (2013) Interlocking feedback loops govern the dynamic behavior of the floral transition in Arabidopsis. *Plant Cell*, vol. 25, pp. 820–33.

[235] E. Ando, M. Ohnishi, Y. Wang, T. Matsushita, A. Watanabe, Y. Hayashi, M. Fujii, J. F. Ma, S.-I. Inoue, T. Kinoshita (2013) TWIN SISTER OF FT, GIGANTEA, and CONSTANS have a positive but indirect effect on blue light-induced stomatal opening in *Arabidopsis thaliana*. *Plant physiol.*, vol. 162, pp. 1529-38.

[236] A. James, J. Monreal, G. Nimmo, C. Kelly (2008) The circadian clock in Arabidopsis roots is a simplified slave version of the clock in shoots. *Science*, vol. 322, pp. 1832–5.

[237] B. Wenden, D. L. K. Toner, S. K. Hodge, R. Grima, A. J. Millar (2012) Spontaneous spatiotemporal waves of gene expression from biological clocks in the leaf. *Proc. Nat. Acad. Sci. USA*, vol. 109, pp. 6757–62.

[238] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V Gutschow, J. M. Jacobs, B. Bolival, N. Assad-Garcia, J. I. Glass, M. W. Covert (2012) A whole-cell computational model predicts phenotype from genotype. *Cell*, vol. 150, pp. 389–401.

[239] T. Clack, A. Shokry, M. Moffet, P. Liu, M. Faul, R. A. Sharrock (2009) Obligate heterodimerization of Arabidopsis Phytochromes C and E and interaction with the PIF3 basic Helix-Loop-Helix transcription factor. *Plant Cell*, vol. 21, pp. 786–99.

[240] M. Bai, J. Shang, E. Oh, M. Fan, Y. Bai (2012) Brassinosteroid, gibberellin and phytochrome impinge on a common transcription module in Arabidopsis. *Nat. Cell Biol.*, vol. 14, pp. 810–7.